Organize

Overview

Thinking back to our real-world scenario, what might we need to do next if we were processing these documents by hand?

We probably don't want to throw all of our invoices into one giant file and call it a day, right? (Well, we might *want* to, but we have to at least pretend like we're responsible!)

If we were ultimately going to store these documents in a file cabinet somewhere, we would need to separate them using file folders so that one invoice doesn't get mixed up with another. We would also need to have some kind of label on the folders so we can easily identify what the folder contains and what information it has.

We have to do the same things with Grooper, and these are done in the Organize phase.

In this phase, we will:

- create a Separation Profile ,

- use the new profile to separate our invoices into individual folders,

- learn about classification and content types,

- train Grooper to recognize our documents so it knows what information to get from them, and

- update our Batch Process to do all of this automatically in the future.

Let's get organized!

Separation

Before we can expect to get any kind of information from these invoices, we need to sort them. While we could get away with doing this manually, Grooper most certainly provides us with a better way.
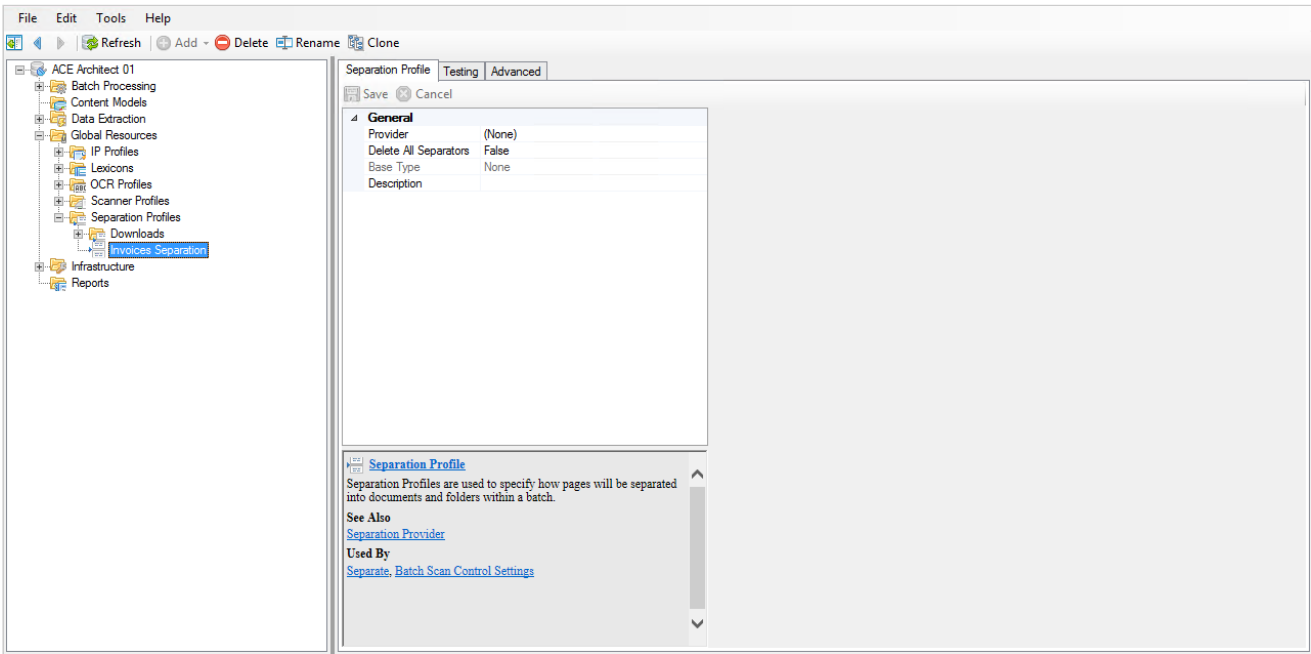
Creating a Separation Profile

> Step 1

1. Navigate to (root) > Global Resources > Separation Profiles
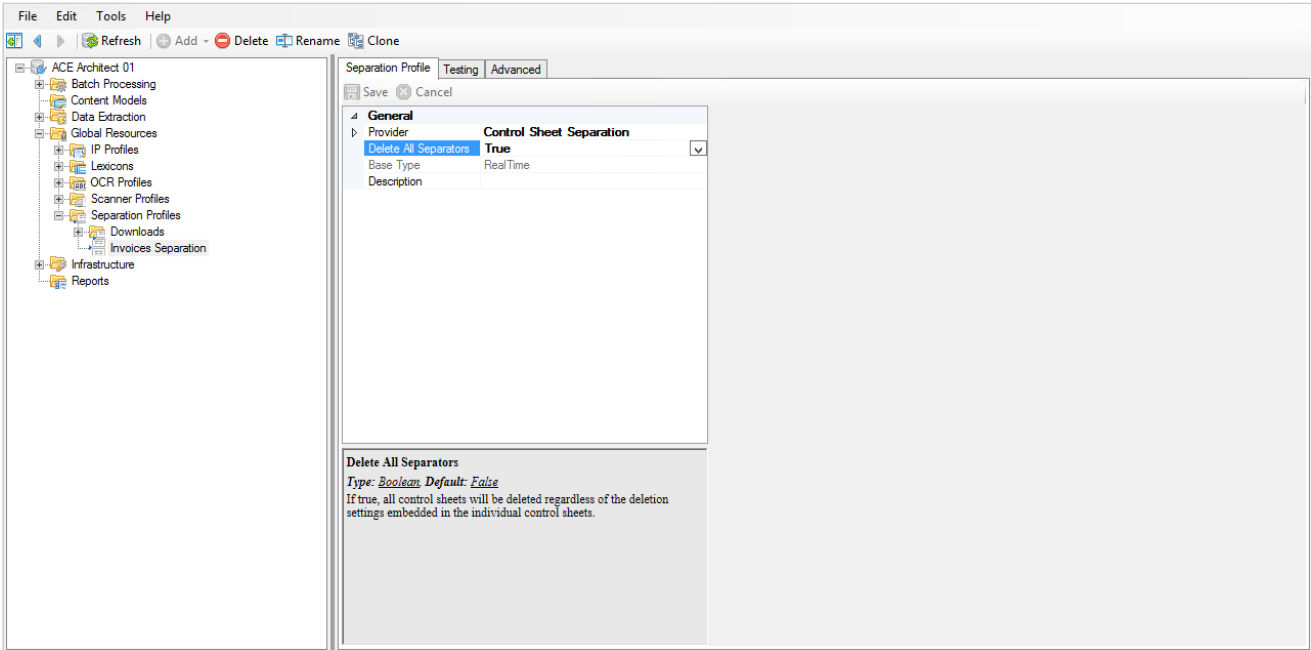
2. Create a new separation profile

3. Give it a name, like Invoices Separation

Step 2

Change the following properties:

1. Set Provider to Control Sheet Separation

2. Set Delete All Separators to True

Updating the production batch

Step 1

1. Navigate to (root) > Batch Processing > Processes > Working > Invoices Process

2. Add a new step

3. Set the Activity Type property of the new step to Separate

4. Set the Separation Profile to the new Invoices Separation profile we created

5. Save and Publish.
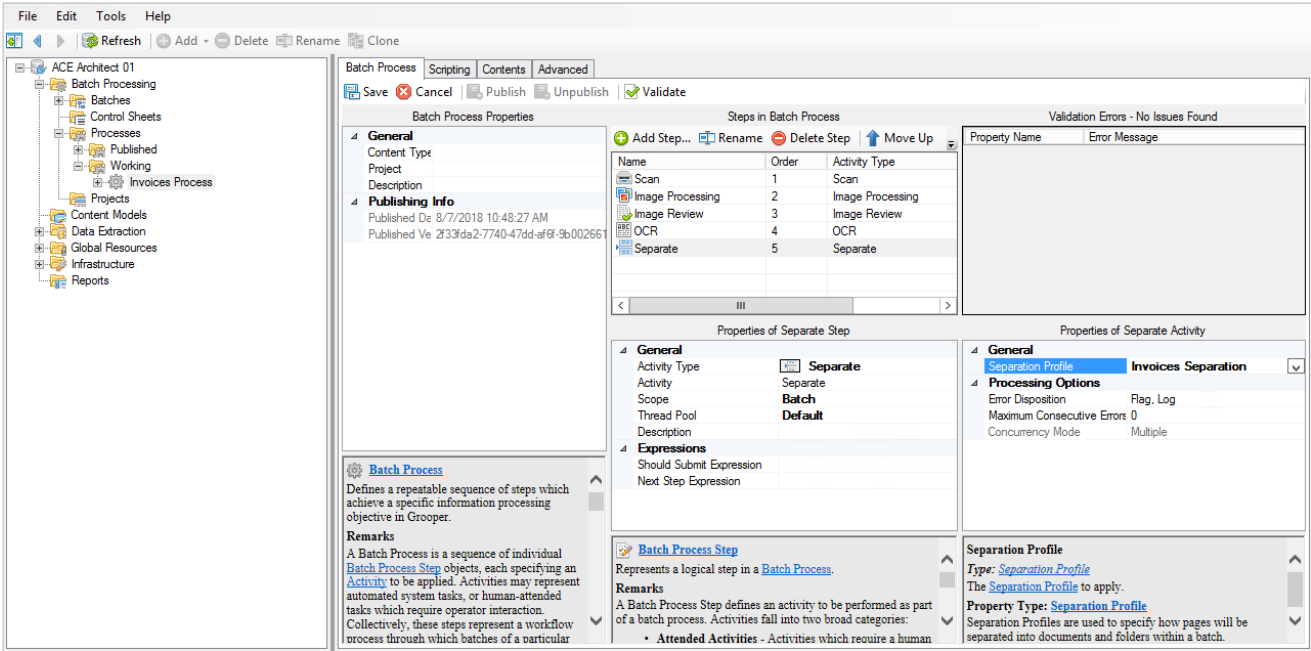
Step 2

1. Navigate to (root) > Batch Processing > Batches > Production > Invoices Process.

2. Make sure the batch is paused

3. Right click on the Invoices Batch and select Update Process….

4. Set Target Step to Separate.

5. Click Execute.
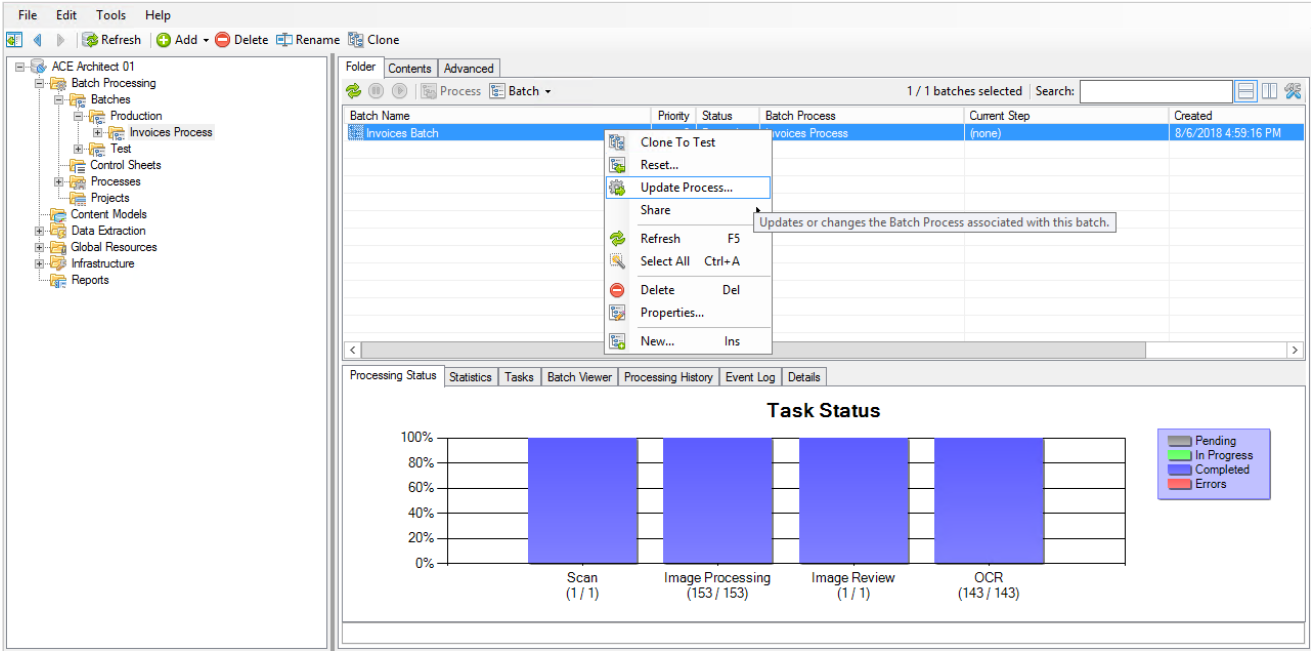
> Step 3

1. Resume the batch

2. Click Process
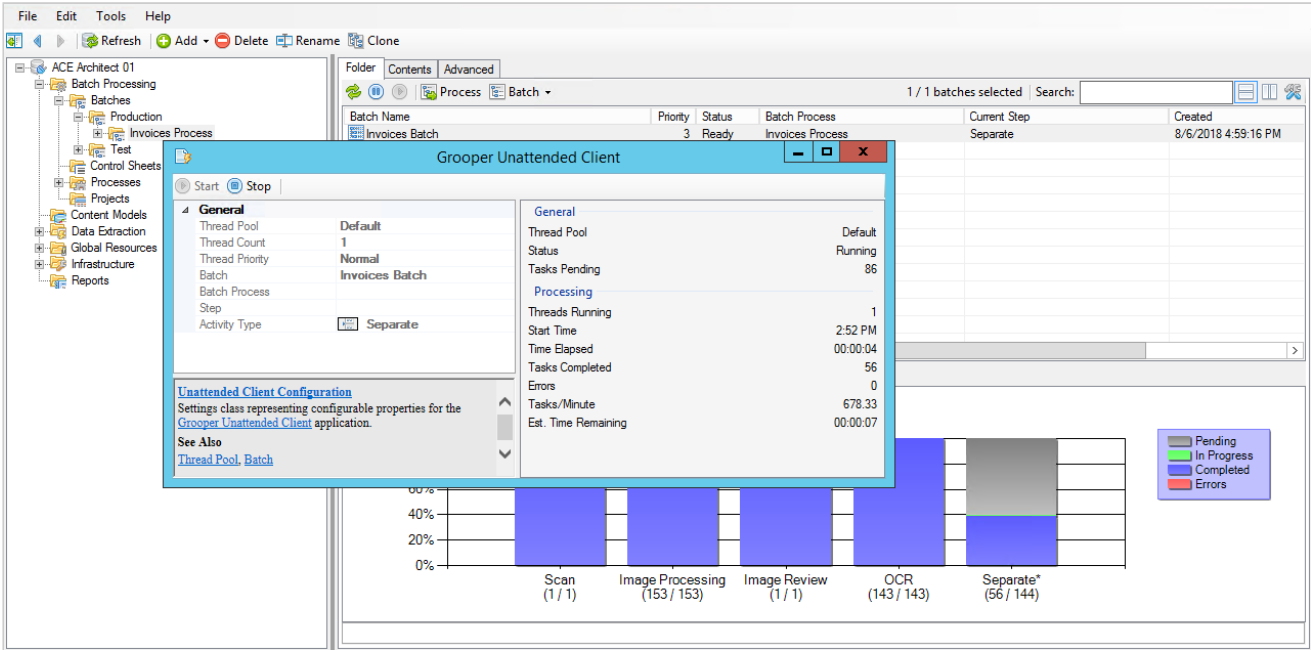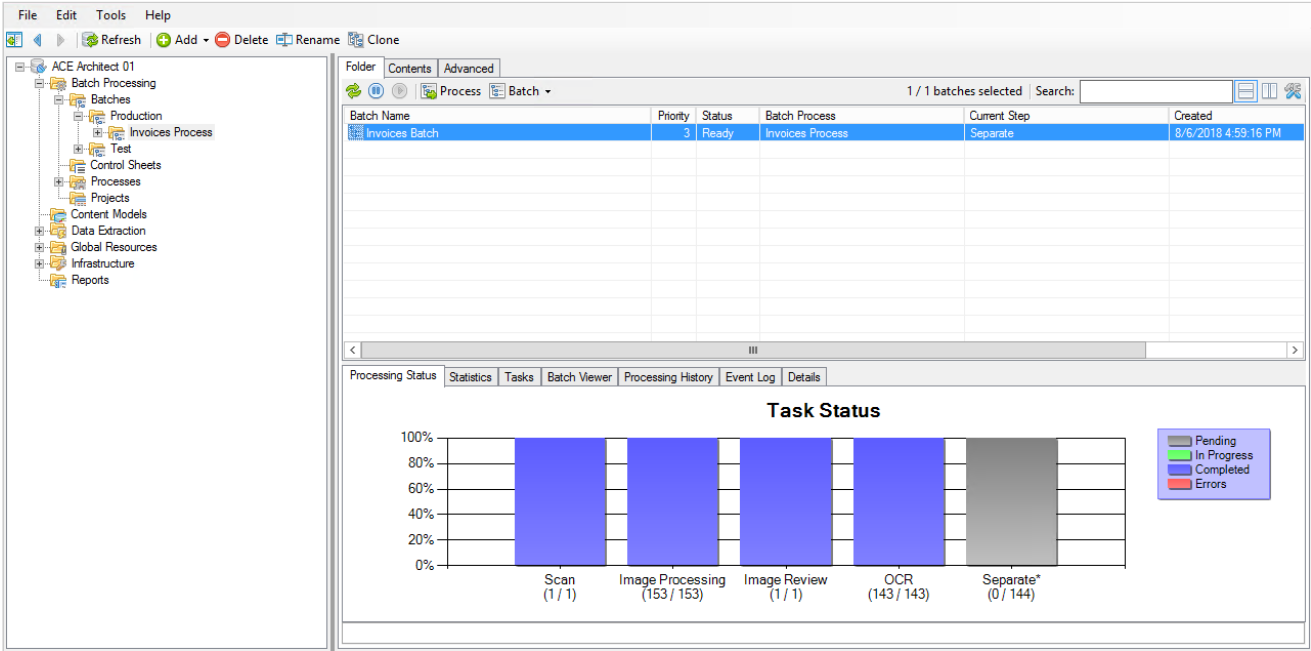
Classification
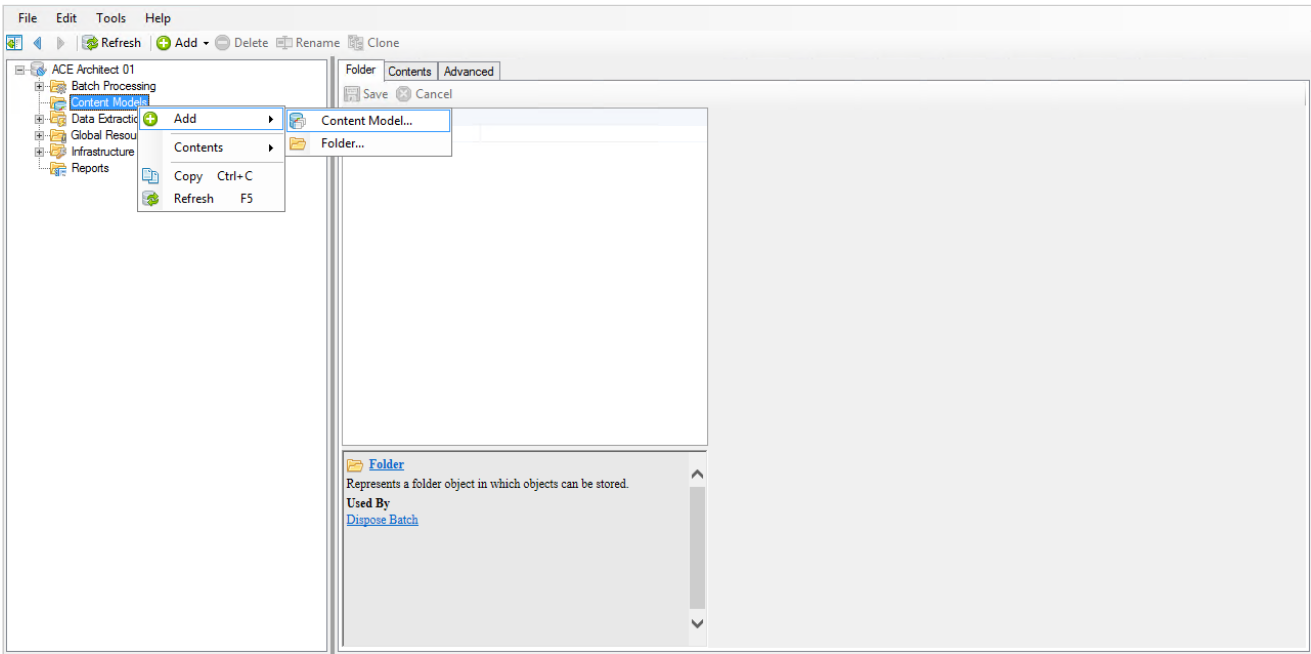
Setting up the Content Model

> Step 1

1. Navigate to (root) > Content Models.

2. Right click on Content Models and select Add > Content Model….

3. Name it Invoices .

1. Right click on the new Invoices Content Model and select Contents > Add Multiple Items....

2. In the Contents • Add Multiple Items window, set Item Type to Document Type

3. Edit the Item Names property and type out the names of each vendor

   - Acme

   - Enid

   - Express

   - Spartan

   - Standard

4. Click Execute to complete
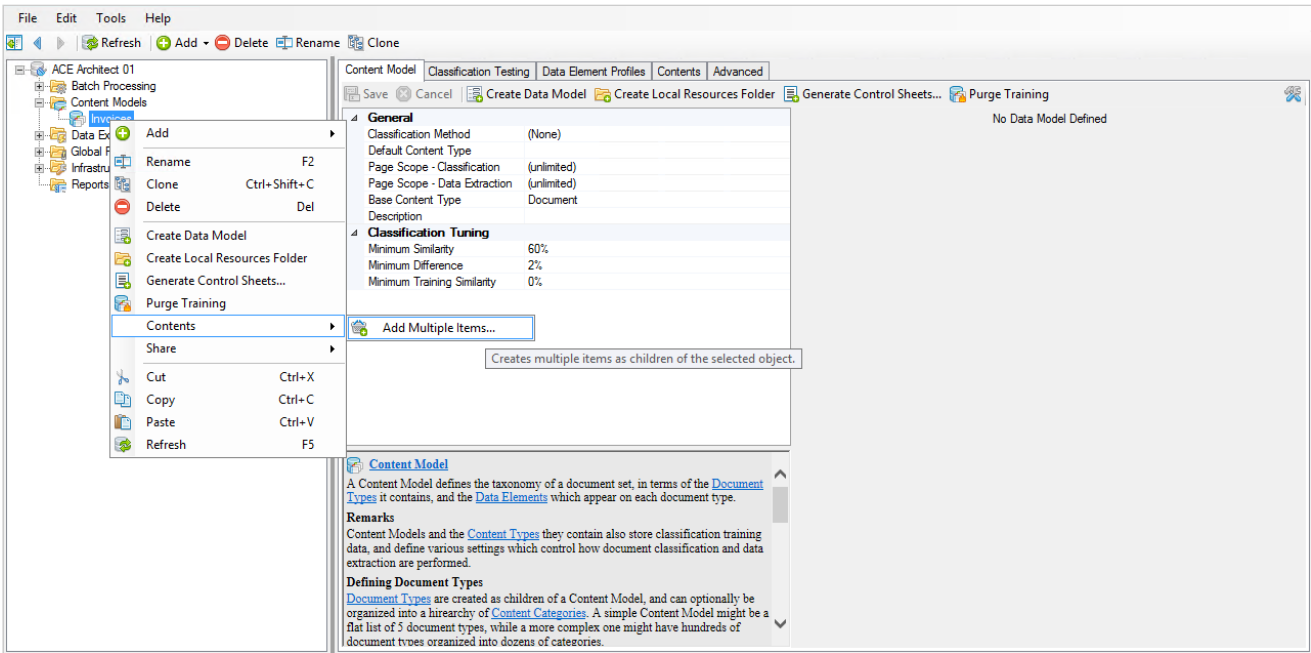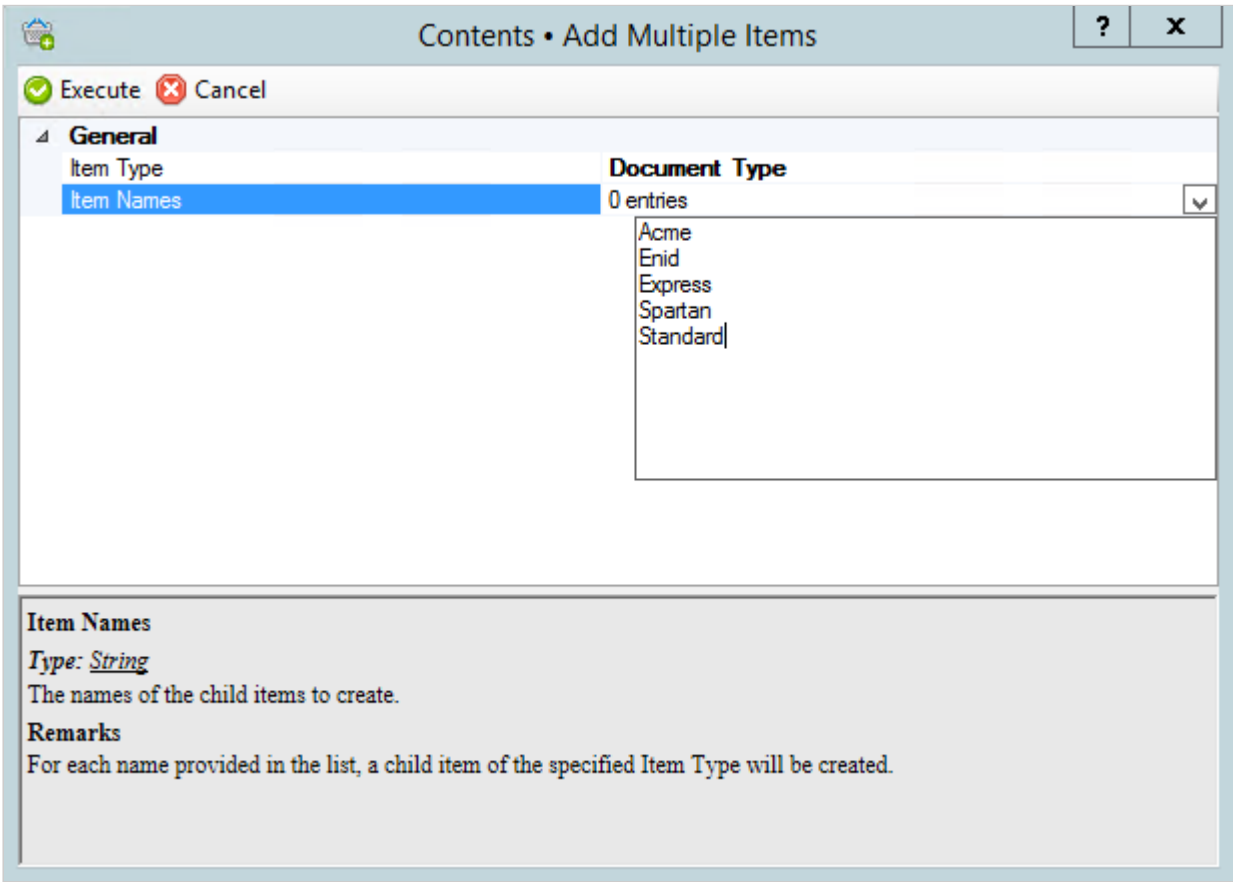
Re-cloning the production batch

> Step 1

1. Navigate to (root) > Batch Processing > Batches > Test > Production Clones > Invoices Process > Invoices Batch

2. Right click on Invoices Batch and select Delete

> Step 2

1. Navigate to (root) > Batch Processing > Batches > Production > Invoices Process > Invoices Batch

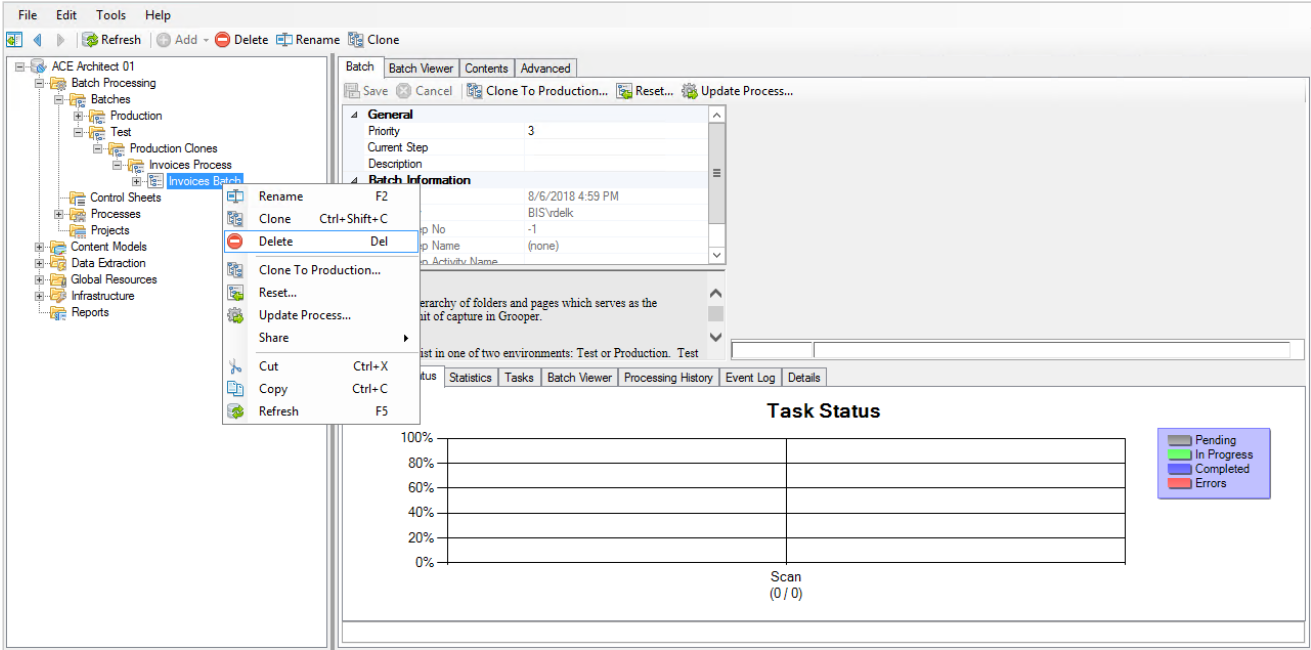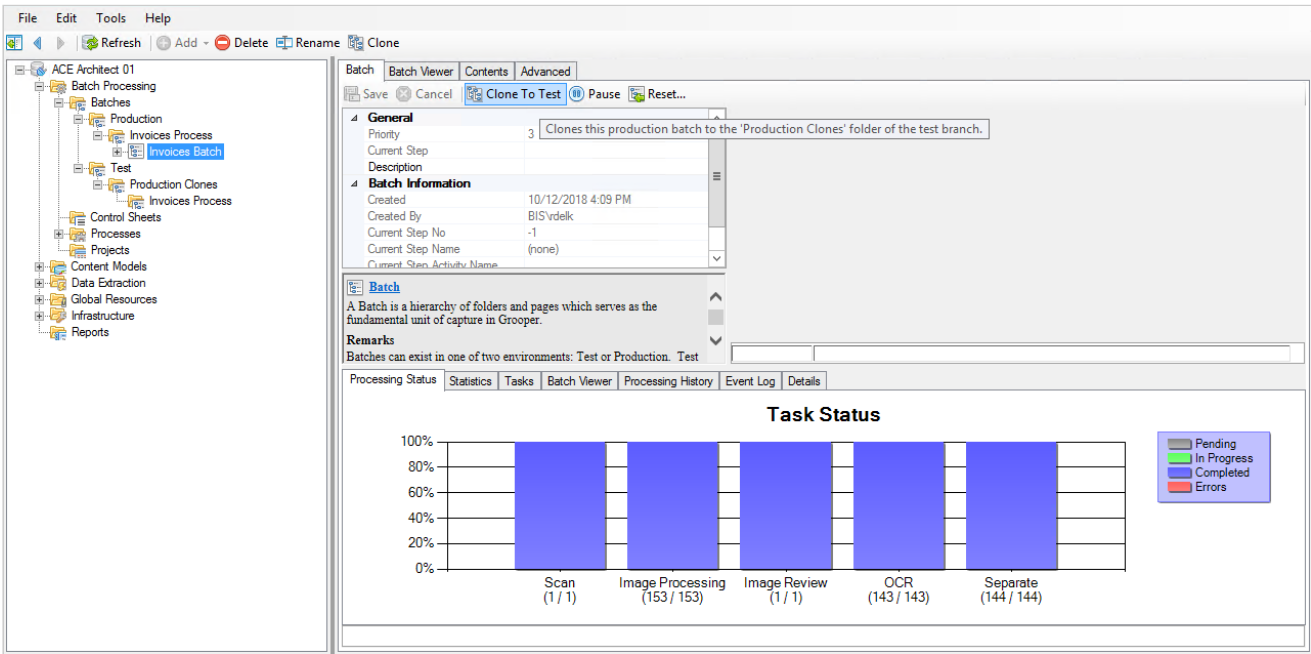2. Click Clone To Test

Setting the classification method

> Step 1

1. Navigate to (root) > Content Models > Invoices

2. Right click on Invoices and select Create Data Model

3. Right click on Invoices and select Create Local Resources Folder
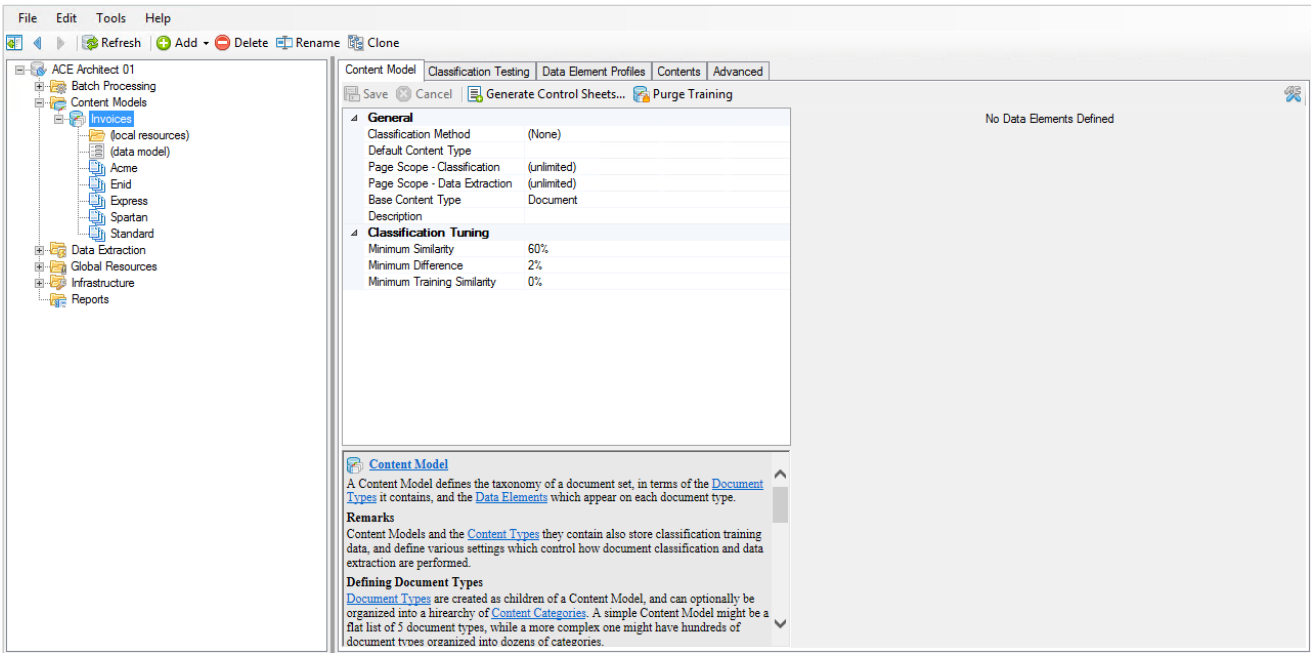
> Step 2

1. On the Invoices Content Model in the General section, set the Classification Method property to Lexical

2. Still in the General section, navigate to Classification Method > Text Feature Extractor > Type

3. Set the Type property to Reference

4. Set the Referenced Extractor property to the Data Extraction > Downloads > Features > Words (Stemmed) Data Type
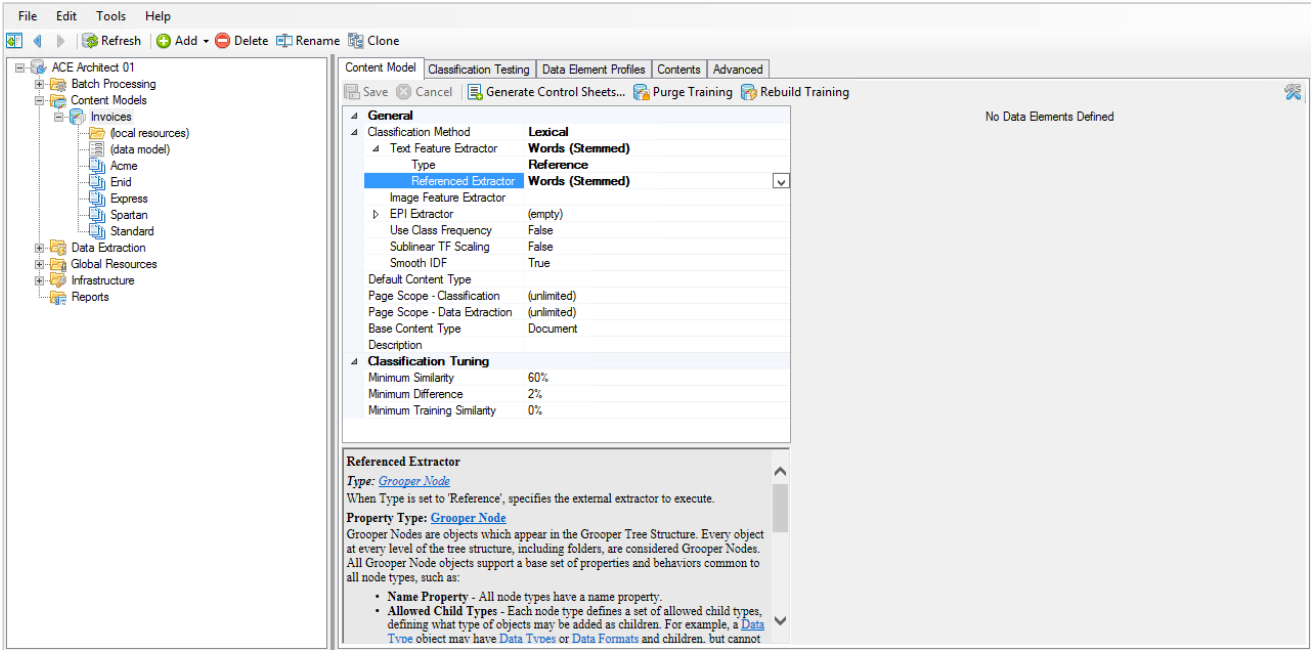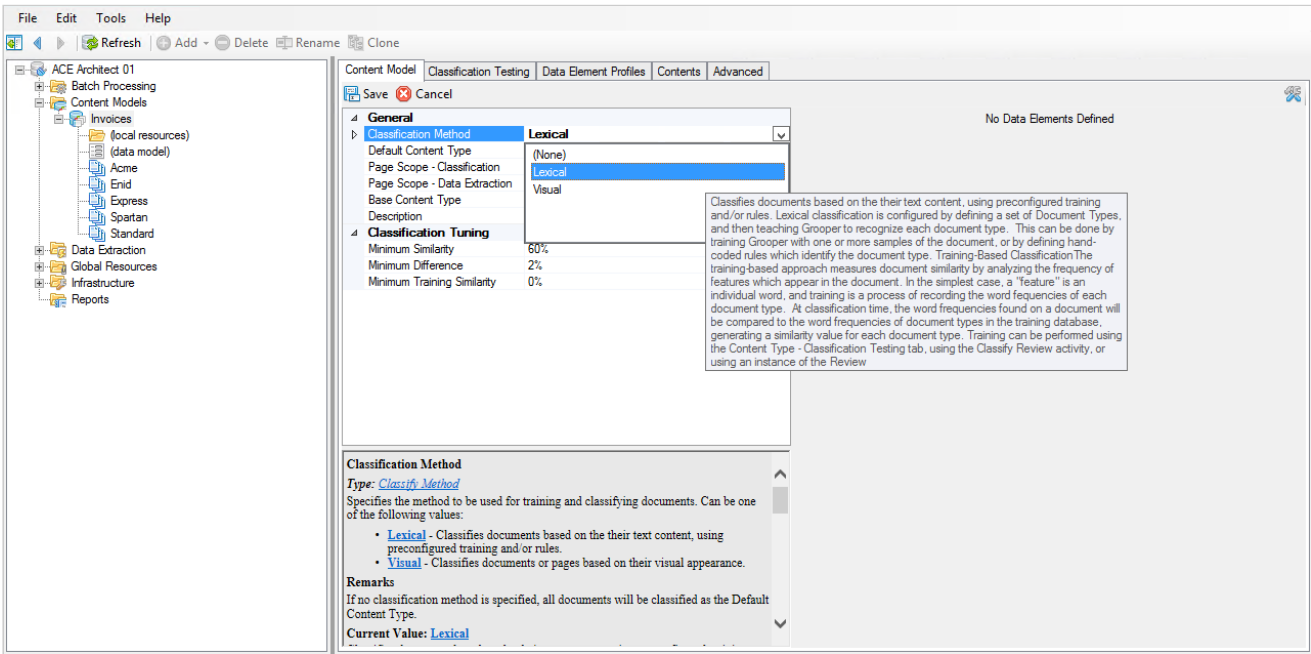
Training and testing classification

> Step 1

1. Click on the Classification Testing tab

2. Select Folder (1)

3. In the toolbar, click on the Train Document button

4. Select the Acme Document Type from the list

5. In the toolbar, click on the Classify Batch button
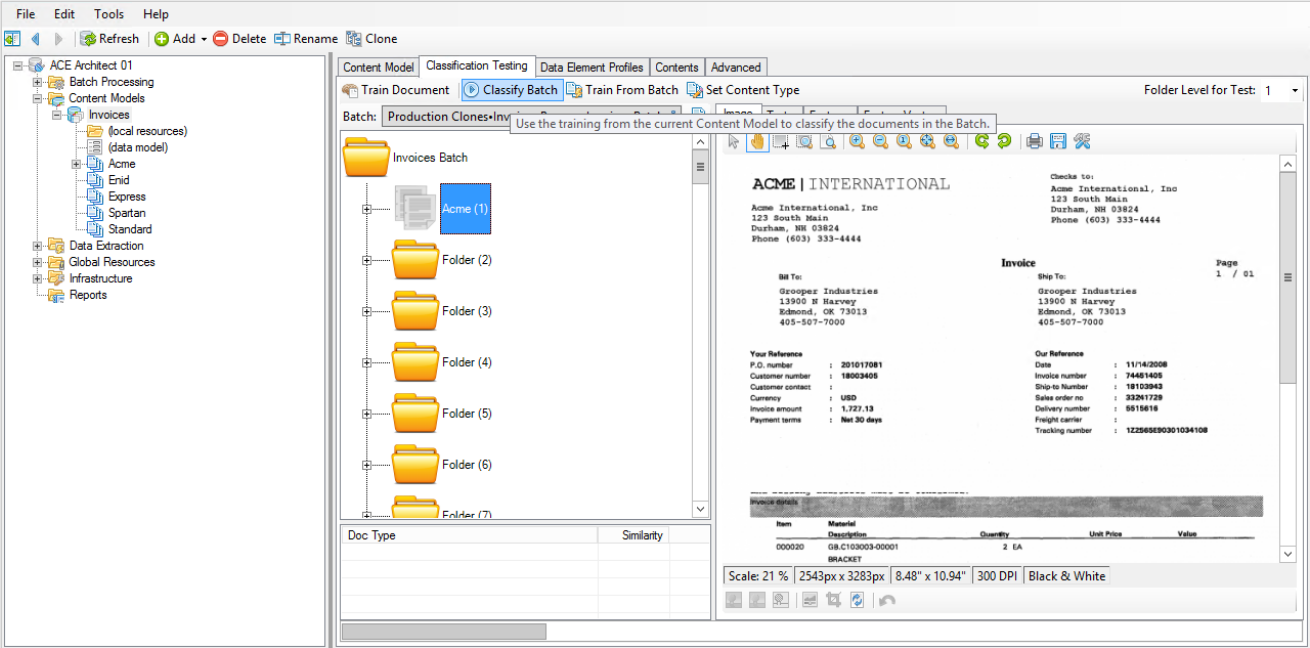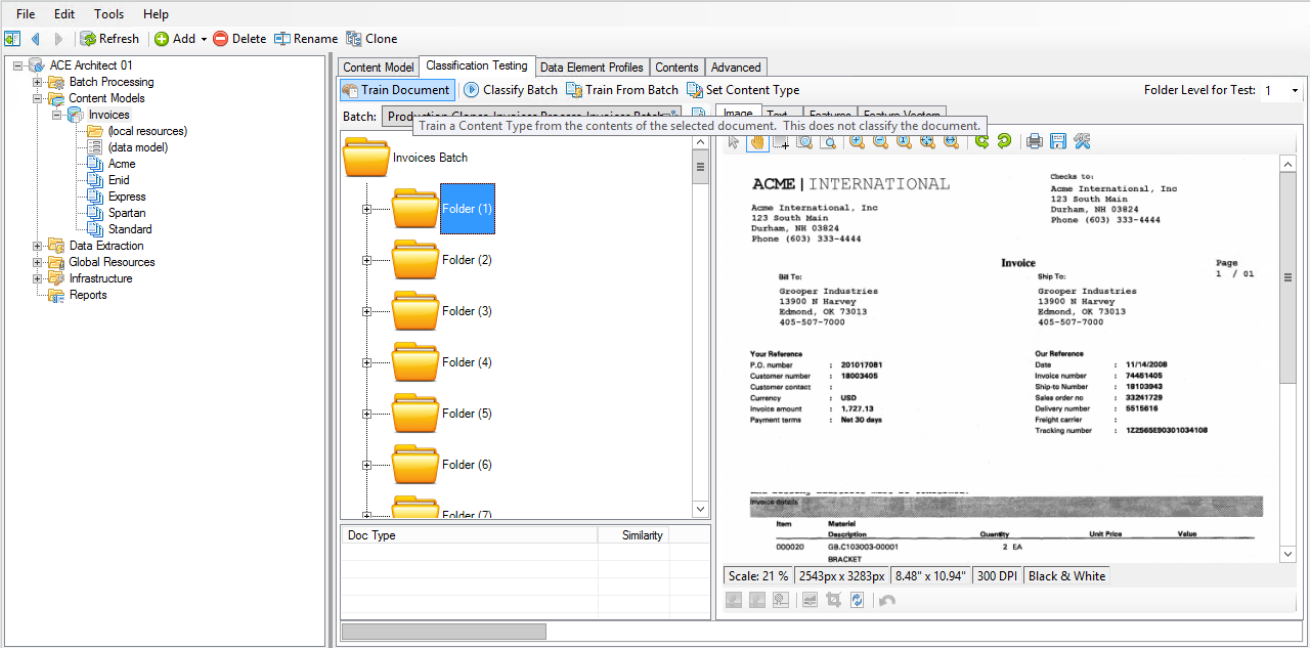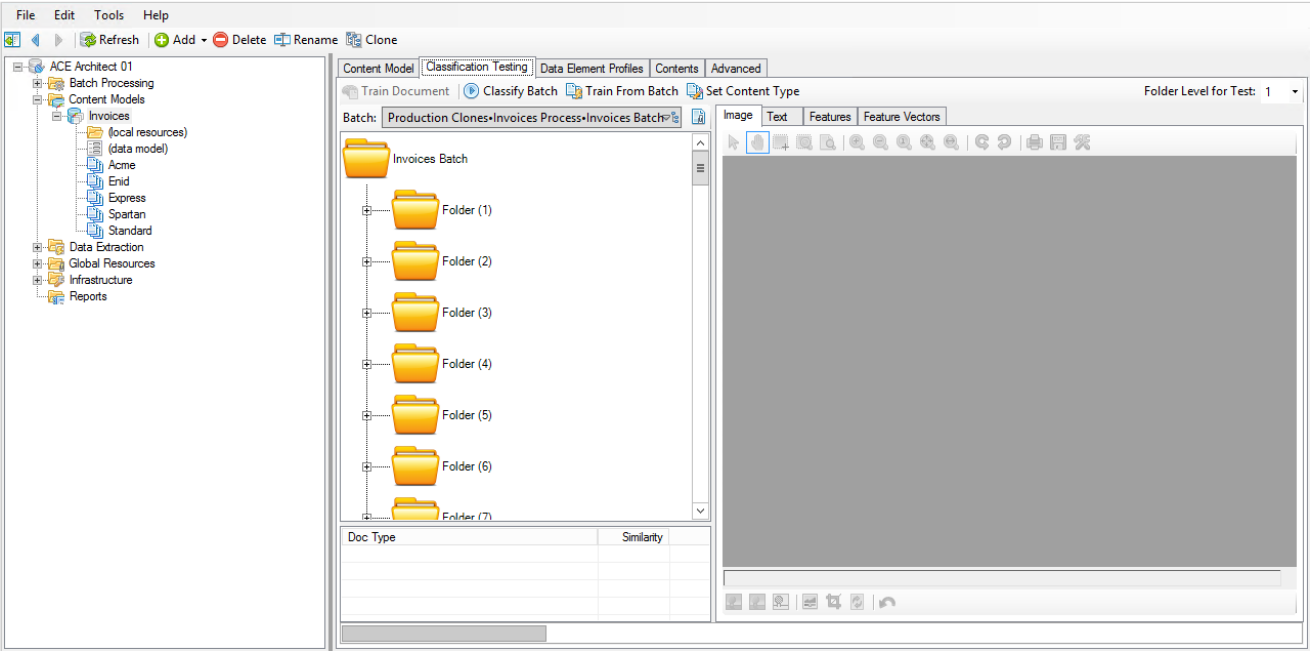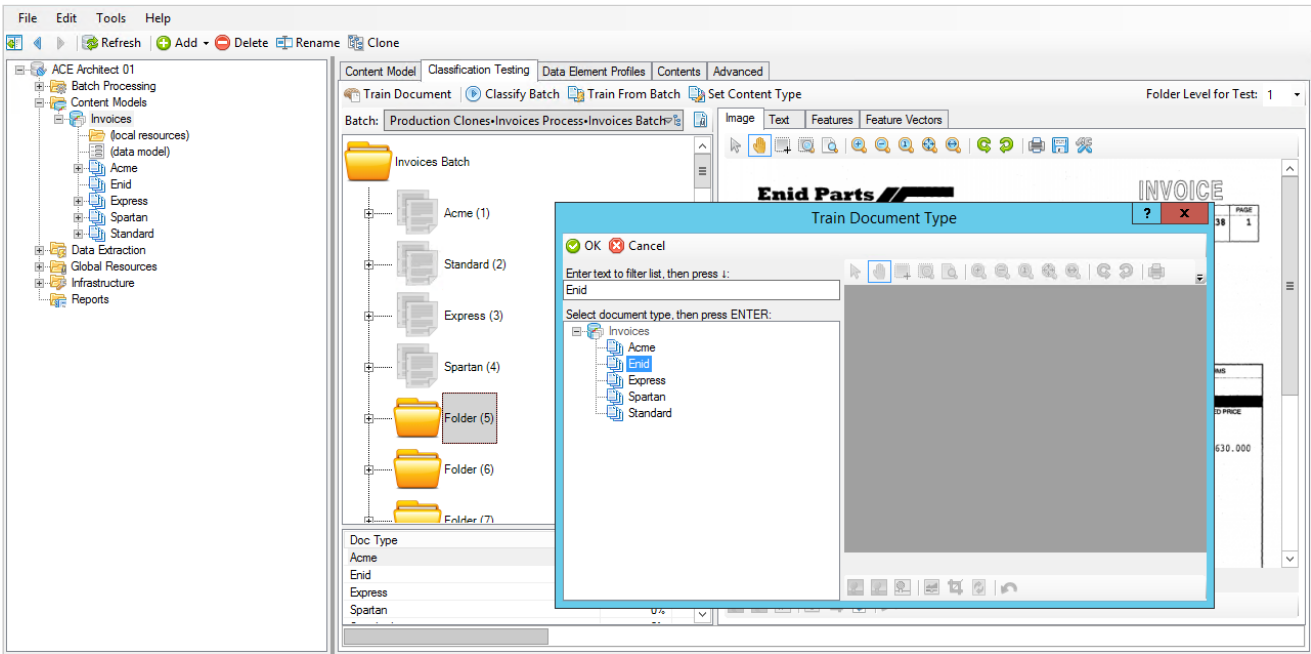
1. Train the other four vendors:

   - Folder (2) ▭ Standard (2)

   - Folder (3) ▭ Express (3)

   - Folder (4) ▭ Spartan (4)

   - Folder (5) ▭ Enid (5)

2. In the toolbar, click on the Classify Batch button
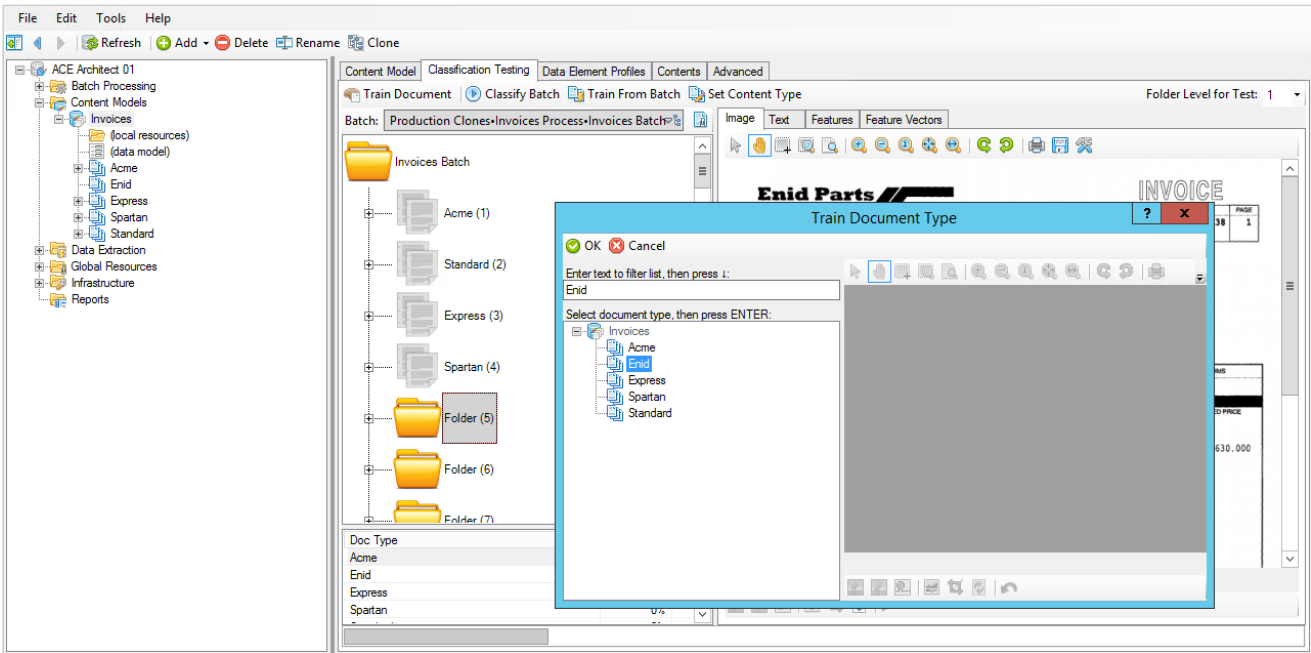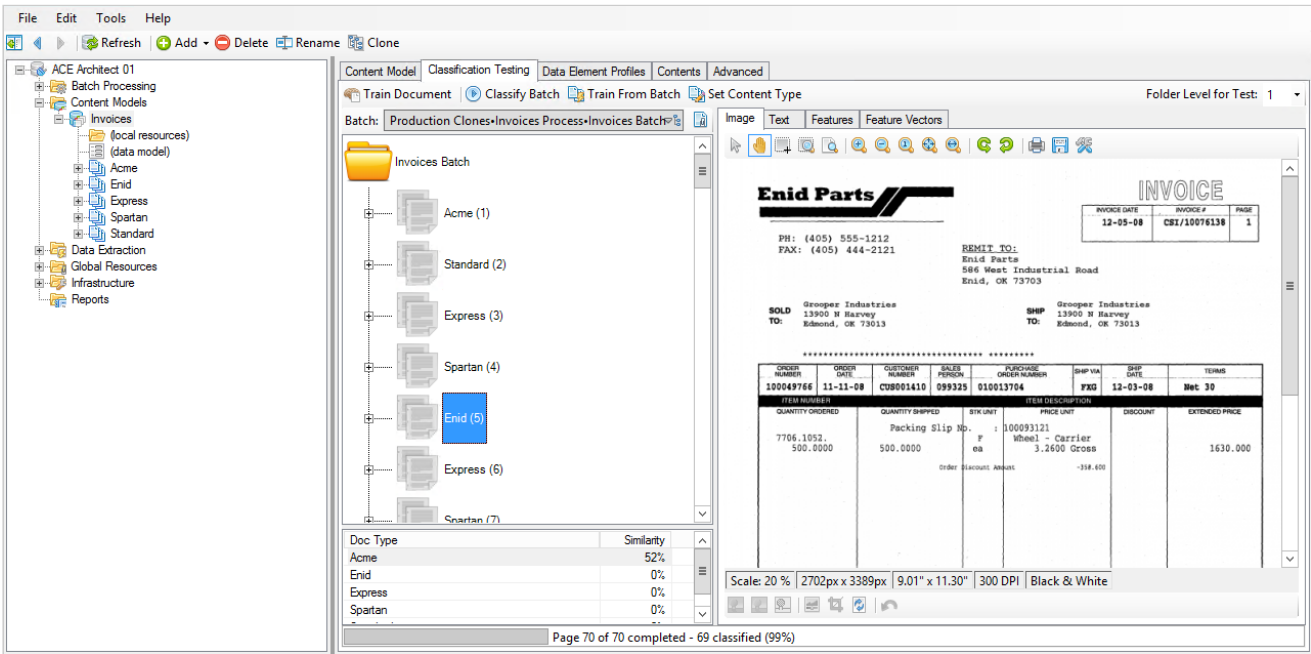
Classify Review

Adding a Classify Review step

> **Step 1**

1. Navigate to (root) > Batch Processing > Processes > Working > Invoices Process

2. Add a new step

3. Set the Activity Type property of the new step to Classify

4. Set the Content Model Scope to our Invoices Content Model
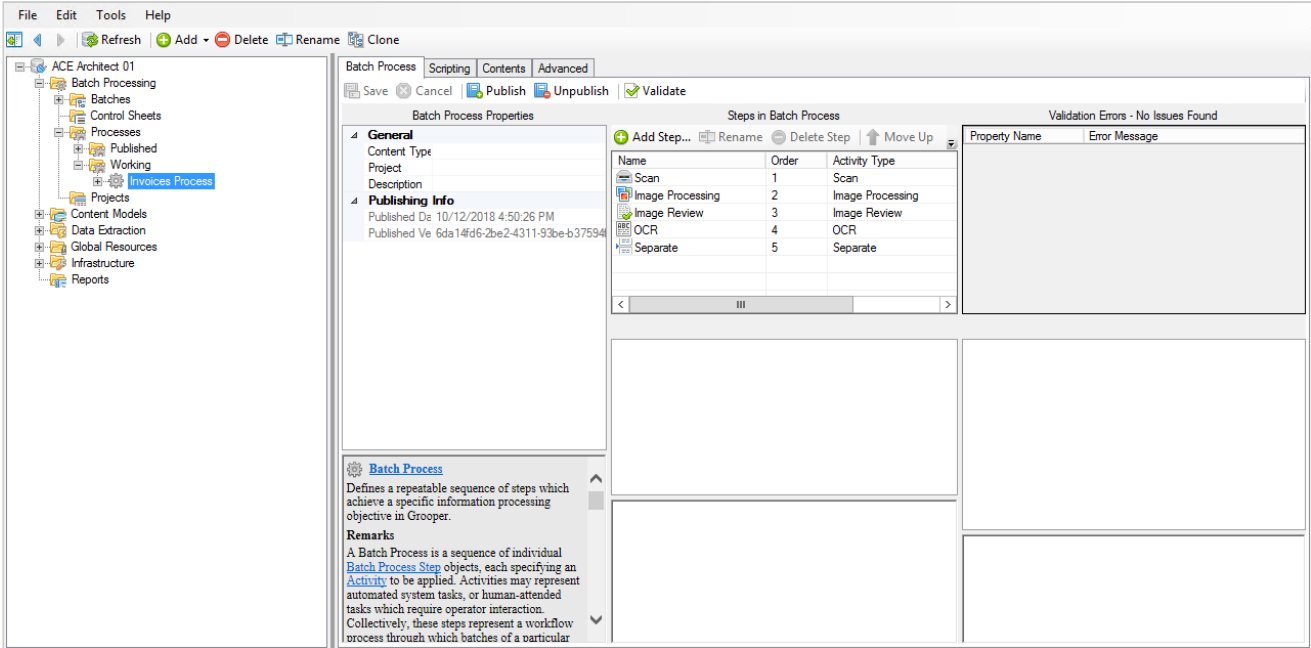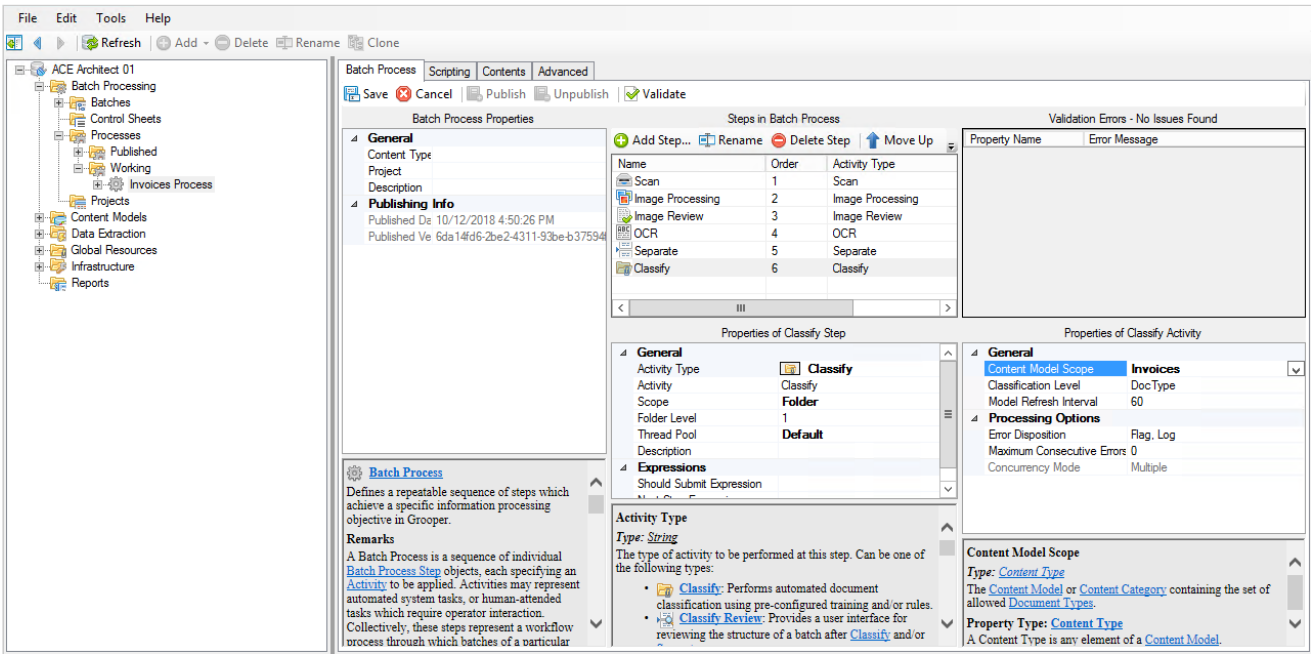
1. Add a new step

2. Set the Activity Type property of the new step to Classify Review

3. Set the Classification Viewer Settings > Content Model Scope to our Invoices Content Model

4. Save and Publish.
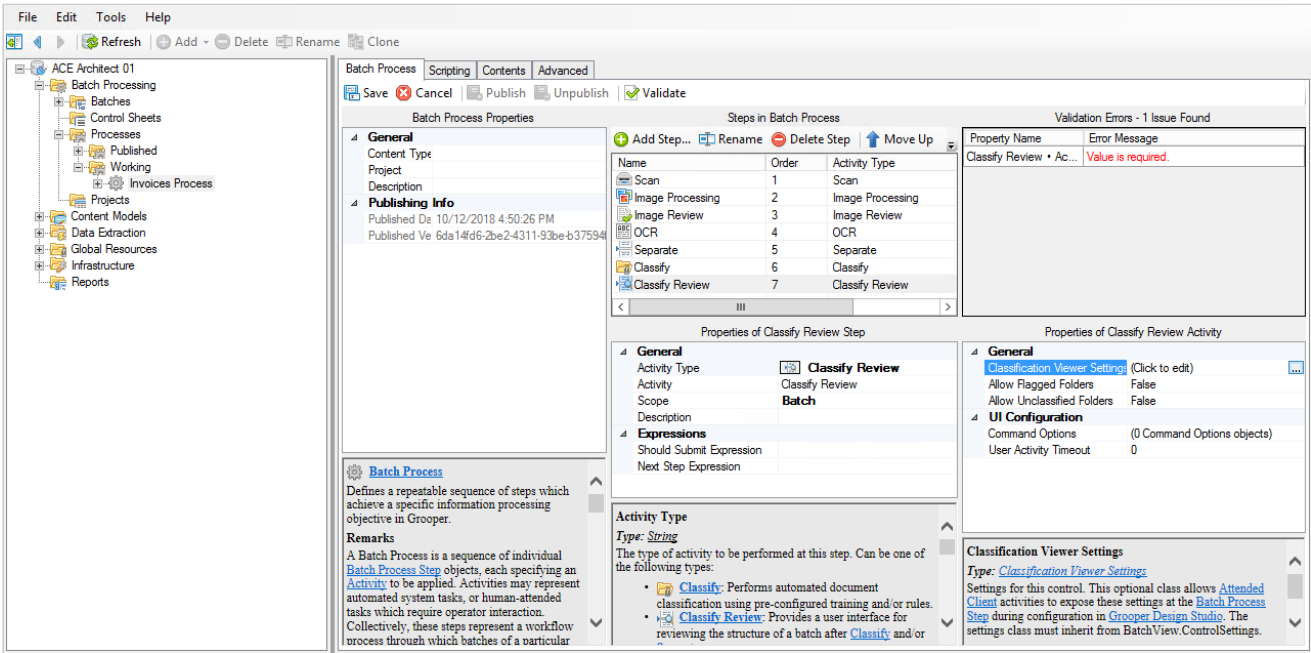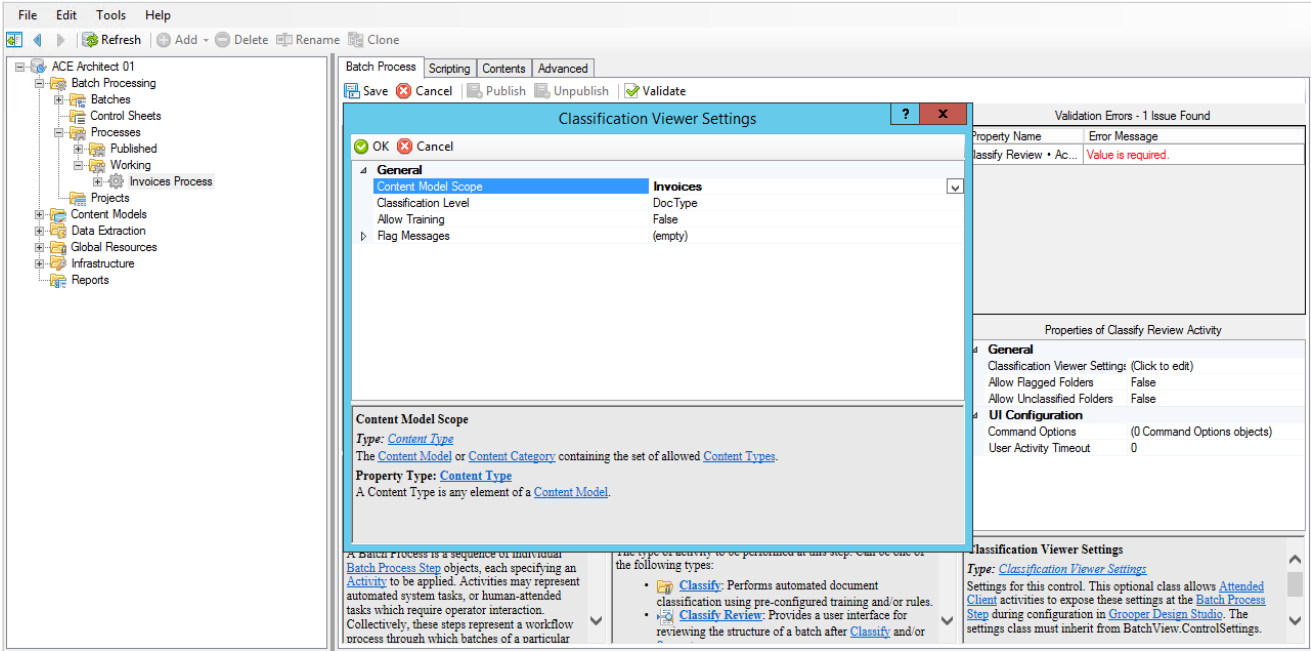
Updating the production batch

> **Step 1**

1. Navigate to (root) > Batch Processing > Batches > Production > Invoices Process

2. Make sure the batch is paused

3. Right click on the Invoices Batch and select Update Process….

4. Set Target Step to Classify.

5. Click Execute.
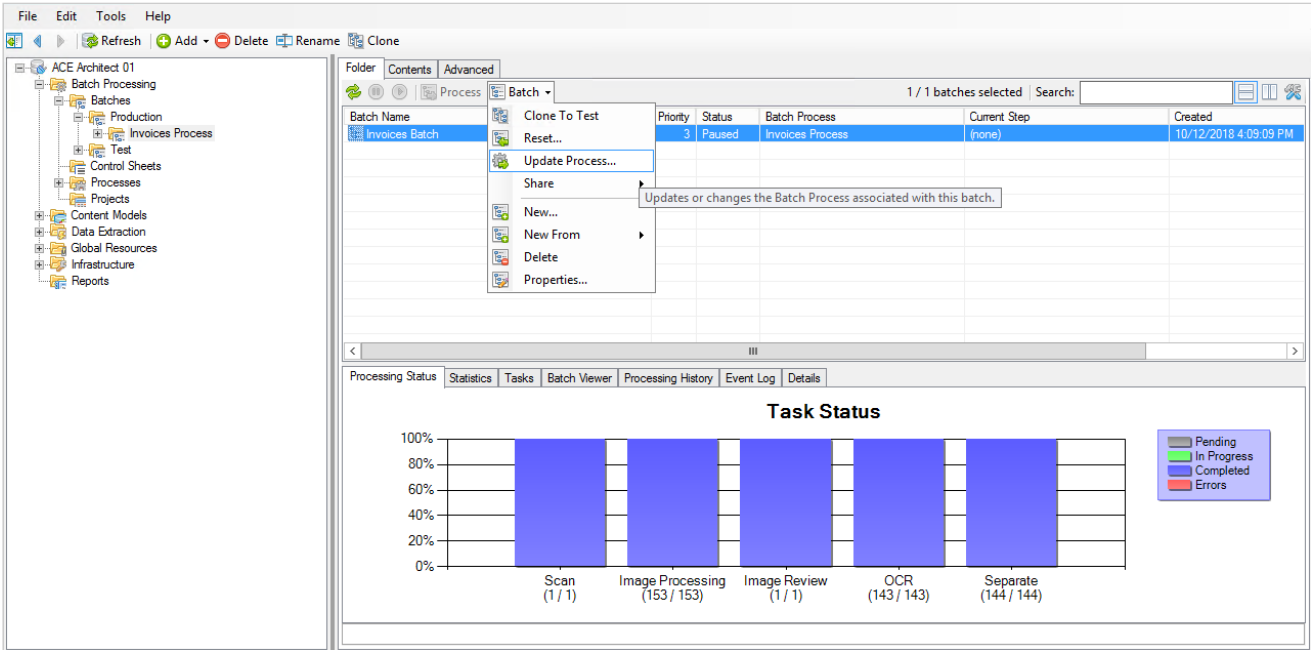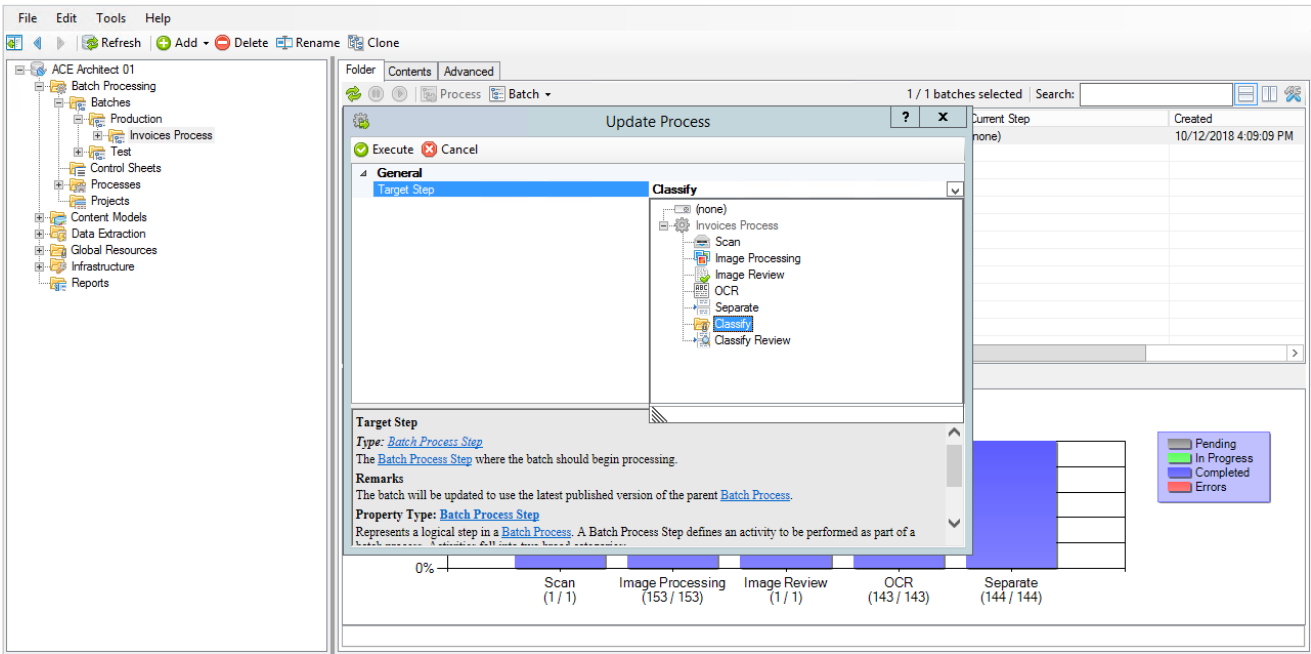
1. Resume the batch
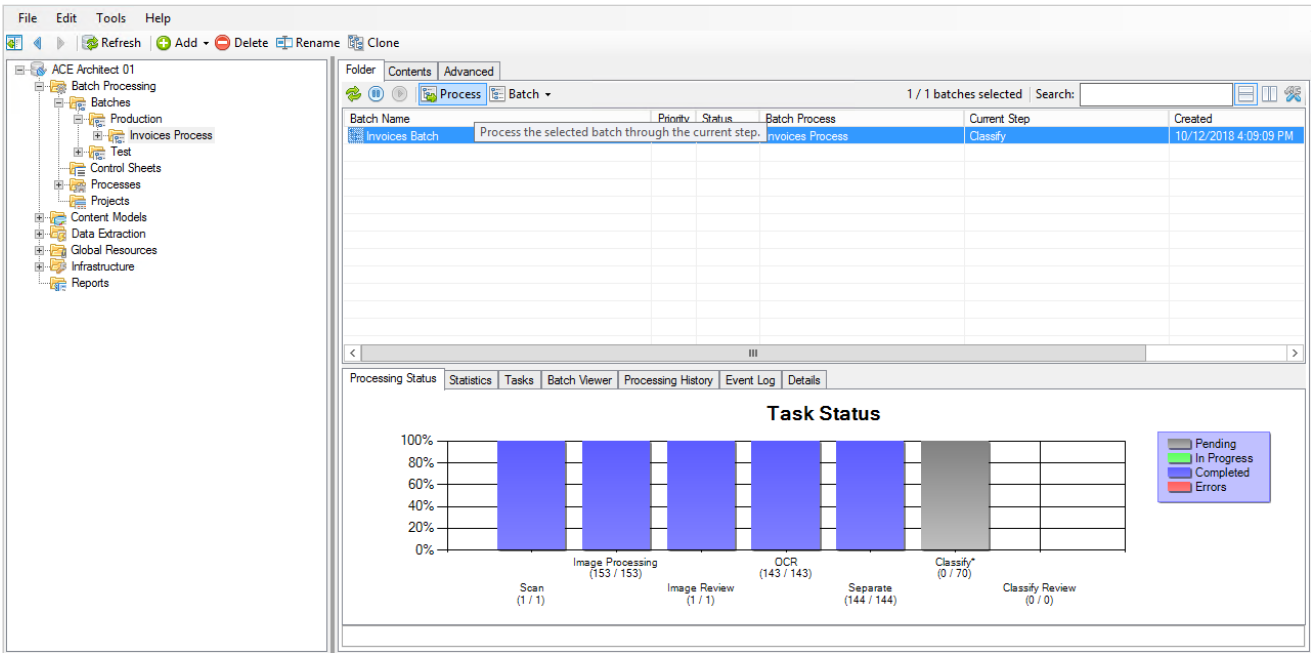
2. Click Process

Reviewing and correcting classification

1. After the Classify step has completed, click Process to start the Classify Review step.

2. In Grooper Classify Review, right click on Folder (18) and select Assign Content Type

3. Select the Enid Document Type from the list and click OK

4. In the toolbar, click on the Complete Task button to finish this activity

Screenshot 1

Processing Task Batch Help
Start Stop Open Task... Close Task Complete Task New Batch... Settings...

Invoices Batch
- Acme (1)
- Standard (2)
- Express (3)
- Spartan (4)
- Enid (5)

Document Type Candidates	%
Acme	100%
Enid	28%
Express	20%
Standard	19%
Spartan	17%

General
Batch Process Name — Invoices Process
Batch Name — Invoices Batch
Batch Open Time — 00:00:35

ACME | INTERNATIONAL

Checks to:
Acme International, Inc
123 South Main
Durham, NH 03824
Phone (603) 333-4444

Acme International, Inc
123 South Main
Durham, NH 03824
Phone (603) 333-4444

Invoice

Page 1 / 01

Bill To:
Grooper Industries
13900 N Harvey
Edmond, OK 73013
405-507-7000

Ship To:
Grooper Industries
13900 N Harvey
Edmond, OK 73013
405-507-7000

Your Reference
P.O. number : 201017081
Customer number : 18003405
Customer contact :
Currency : USD
Invoice amount : 1,727.13
Payment terms : Net 30 days

Our Reference
Date : 11/14/2008
Invoice number : 74451405
Ship-to Number : 18103943
Sales order no : 33241729
Delivery number : 5515616
Freight carrier :
Tracking number : 1Z2565E90301034108

Scale: 34 % 2543px x 3283px 8.48" x 10.94" 300 DPI Black & White

Screenshot 2

Processing Task Batch Help
Start Stop Open Task... Close Task Complete Task New Batch... Settings...

- Standard (15)
- Acme (16)
- Express (17)
- Folder...
- Acme...
- Express...

Context menu:
Cut	Ctrl+X	
Copy	Ctrl+C	
Paste	Ctrl+V	
Delete	Del	
Append to Previous	Ctrl+P	
Assign Content Type	Ctrl+F	
Flag Item...	Ctrl+I	
Go To Next Issue	Ctrl+N	
Goto Next Flagged Item	Ctrl+G	
New Folder	Ins	
Unclassify	Ctrl+U	
Batch Operations		
Contents		
Send To		
Move Up	Ctrl+Up	
Move Down	Ctrl+Down	
Item Properties...		

Assigns a Content Type to the currently selected folder.

Document Type Candidates
Enid
Acme
Express
Spartan
Standard

General
Batch Process Name
Batch Name
Batch Open Time

Enid Parts
INVOICE
PH: (405) 555-1212
FAX: (405) 444-2121

INVOICE DATE	INVOICE #	PAGE
12-08-08	CSI/10076269	1

REMIT TO:
Enid Parts
586 West Industrial Road
Enid, OK 73703

SOLD TO: Grooper Industries 13900 N Harvey Edmond, OK 73013
SHIP TO: Grooper Industries 13900 N Harvey Edmond, OK 73013

ORDER NUMBER	ORDER DATE	CUSTOMER NUMBER	SALES PERSON	PURCHASE ORDER NUMBER	SHIP VIA	SHIP DATE	TERMS
100050218	12-04-08	CUS001605	099321	201017541	FXG	12-05-08	Net 30

ITEM NUMBER / ITEM DESCRIPTION

QUANTITY ORDERED	QUANTITY SHIPPED	STK UNIT	PRICE UNIT	DISCOUNT	EXTENDED PRICE
	Packing Slip No. : 100093143				
7710.1111.			Housing, Cold Head		
5.0000	4.0000	ea	719.2800 Gross		2877.120
7710.1104.		G	Arm, Latch Cold Head		
4.0000	5.0000	ea	194.4000 Gross		972.000
7710.1140.		E	Cover Rear Hot Head		

Scale: 32 % 2750px x 3437px 9.17" x 11.46" 300 DPI Black & White

Screenshot 3

Processing Task Batch Help
Start Stop Open Task... Close Task Complete Task New Batch... Settings...

- Standard (15)
- Acme (16)
- Express (17)
- Enid (18)
- Acme (19)
- Express (20)

Document Type Candidates	%
Enid	58%
Acme	20%
Express	20%
Spartan	20%
Standard	17%

General
Batch Process Name — Invoices Process
Batch Name — Invoices Batch
Batch Open Time — 00:01:21

Classify Review
Assign Content Type — 1

Enid Parts
INVOICE
PH: (405) 555-1212
FAX: (405) 444-2121

INVOICE DATE	INVOICE #	PAGE
12-08-08	CSI/10076269	1

REMIT TO:
Enid Parts
586 West Industrial Road
Enid, OK 73703

SOLD TO: Grooper Industries 13900 N Harvey Edmond, OK 73013
SHIP TO: Grooper Industries 13900 N Harvey Edmond, OK 73013

ORDER NUMBER	ORDER DATE	CUSTOMER NUMBER	SALES PERSON	PURCHASE ORDER NUMBER	SHIP VIA	SHIP DATE	TERMS
100050218	12-04-08	CUS001605	099321	201017541	FXG	12-05-08	Net 30

ITEM NUMBER / ITEM DESCRIPTION

QUANTITY ORDERED	QUANTITY SHIPPED	STK UNIT	PRICE UNIT	DISCOUNT	EXTENDED PRICE
	Packing Slip No. : 100093143				
7710.1111.			Housing, Cold Head		
5.0000	4.0000	ea	719.2800 Gross		2877.120
7710.1104.		G	Arm, Latch Cold Head		
4.0000	5.0000	ea	194.4000 Gross		972.000
7710.1140.		E	Cover Rear Hot Head		

Scale: 32 % 2750px x 3437px 9.17" x 11.46" 300 DPI Black & White

Refresh   Add ▾   Delete   Rename   Clone

ACE Architect 01
  Batch Processing
    Batches
      Production
        Invoices Process
      Test
    Control Sheets
    Processes
    Projects
  Content Models
  Data Extraction
  Global Resources
  Infrastructure
  Reports

Folder | Contents | Advanced

Process   Batch ▾   1 / 1 batches selected   Search:

Batch Name	Priority	Status	Batch Process	Current Step	Created
Invoices Batch	3	Working	Invoices Process	Classify Review	10/12/2018 4:09:09 PM

Processing Status | Statistics | Tasks | Batch Viewer | Processing History | Event Log | Details

Task Status

100%
80%
60%
40%
20%
0%

Scan (1 / 1)   Image Processing (153 / 153)   Image Review (1 / 1)   OCR (143 / 143)   Separate (144 / 144)   Classify (70 / 70)   Classify Review* (1 / 1)

Legend:
■ Pending
■ In Progress
■ Completed
■ Errors