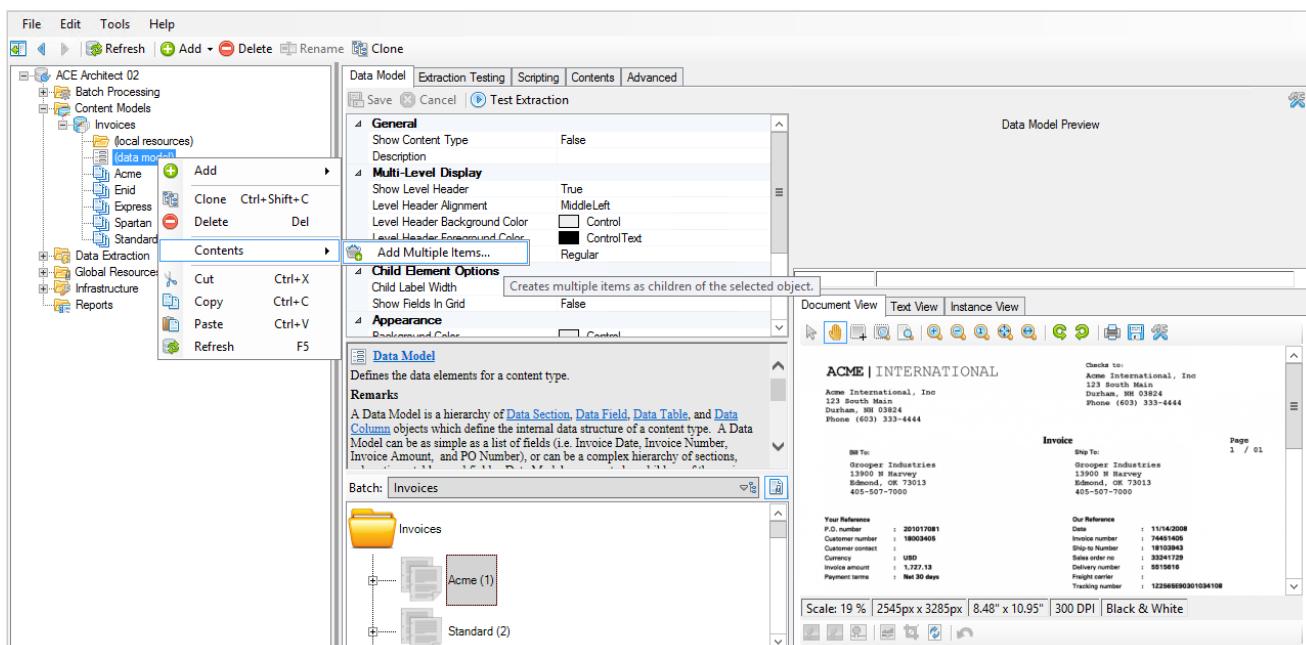


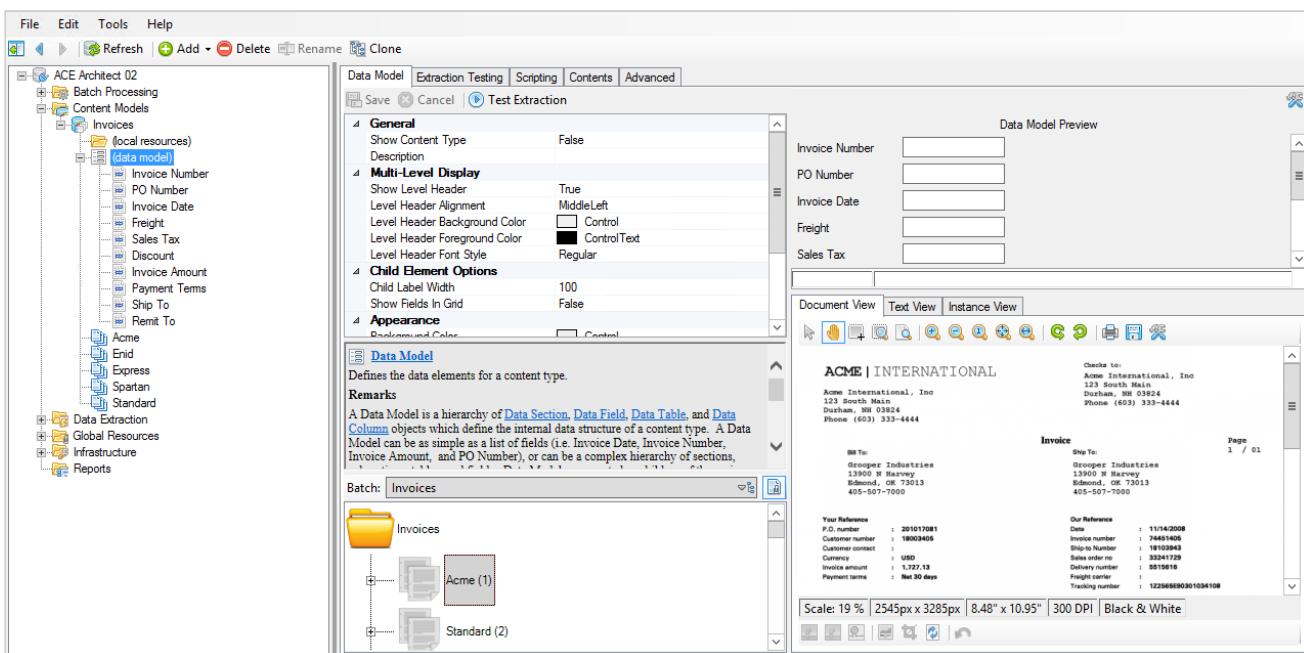
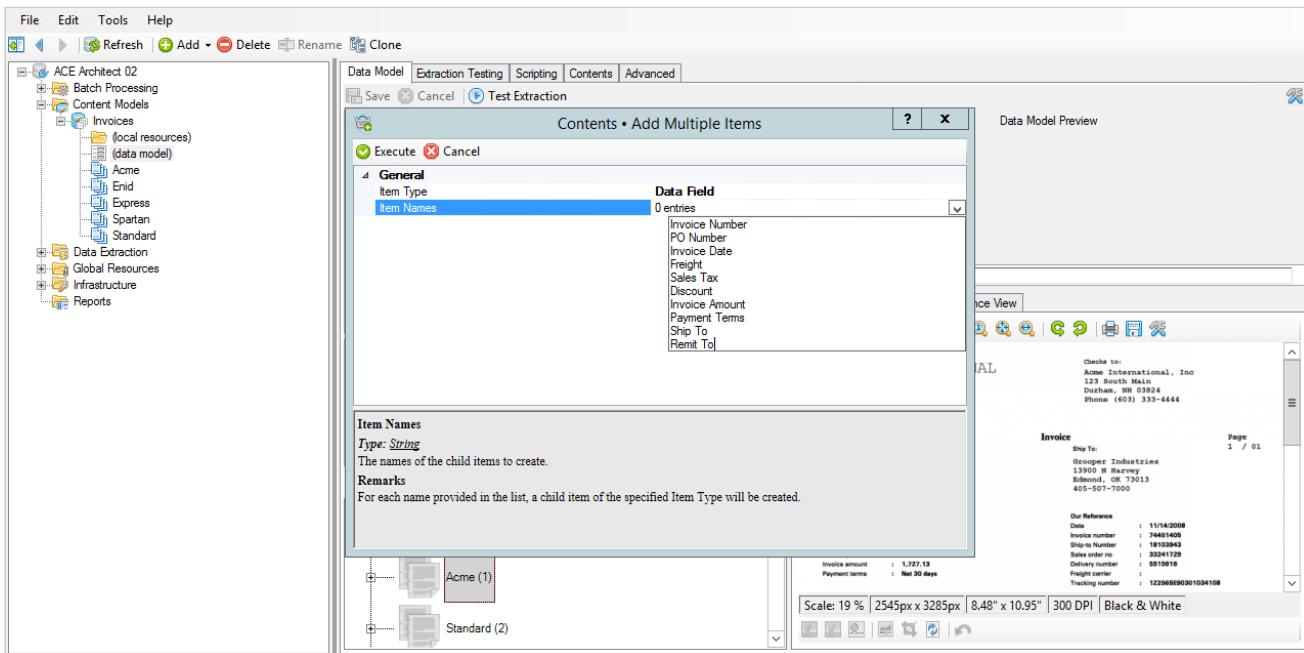
Collect

Invoice Date

> Content Models / Invoices / **(data model)**

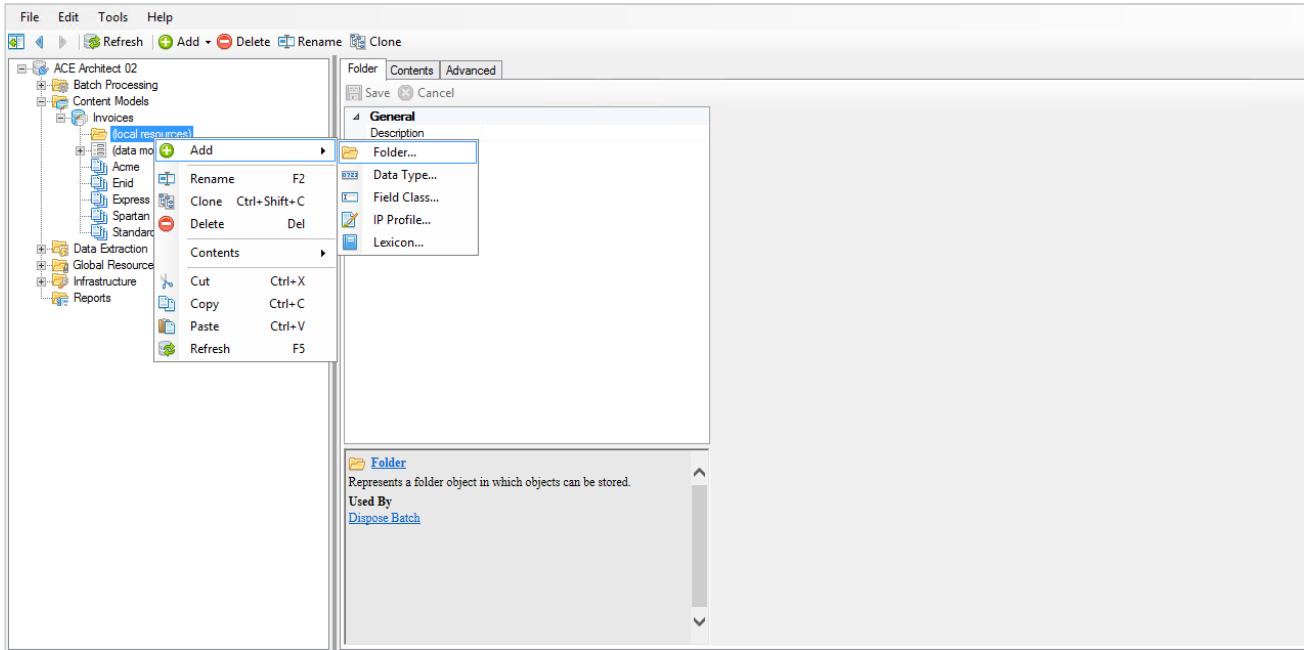
1. Right click and select **Contents > Add Multiple Items...**.
2. In the Contents • Add Multiple Items window, make sure Item Type is set to **Data Field**
3. Edit the Item Names property and type out the names of the following fields
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
4. Click Execute to complete

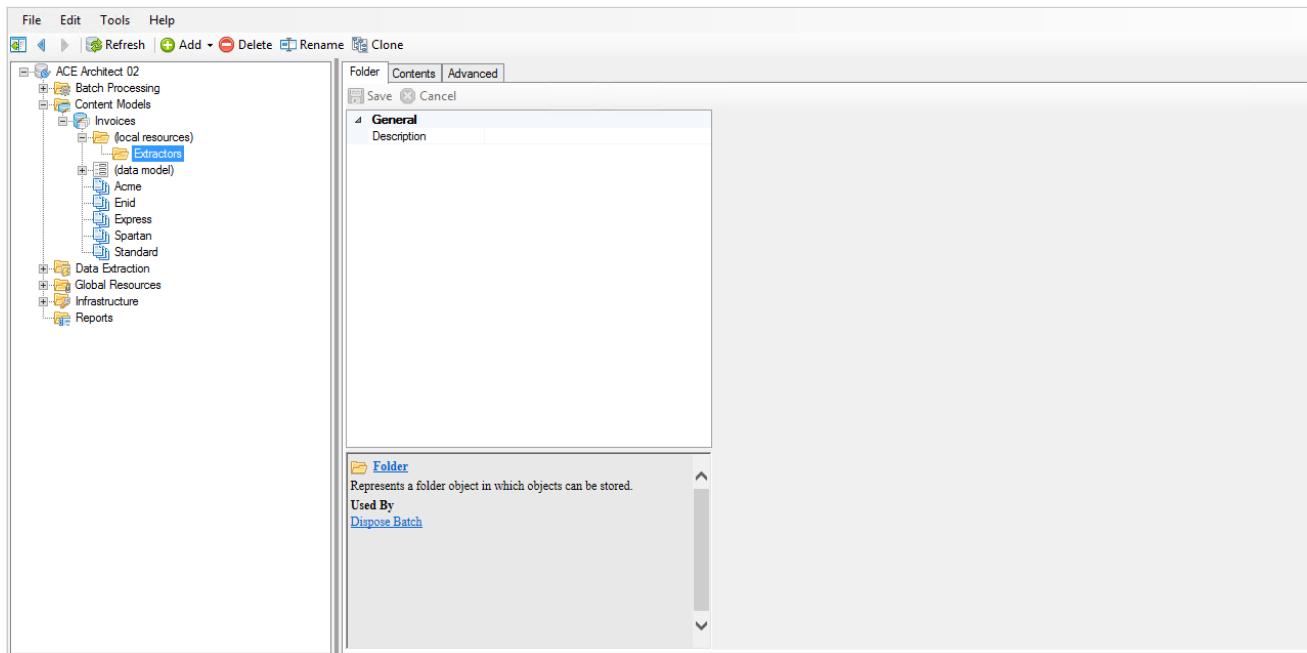




Content Models / Invoices / (local resources)

1. Right click and select Add > Folder...
2. Give it a name, like Extractors

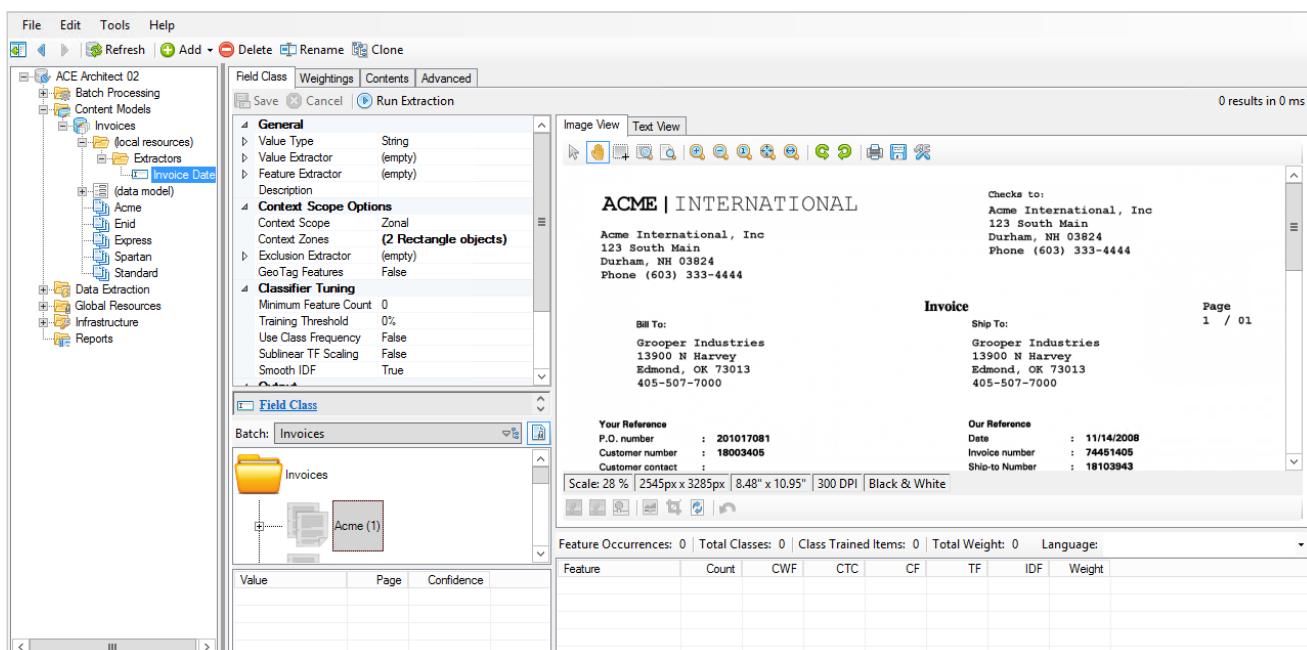
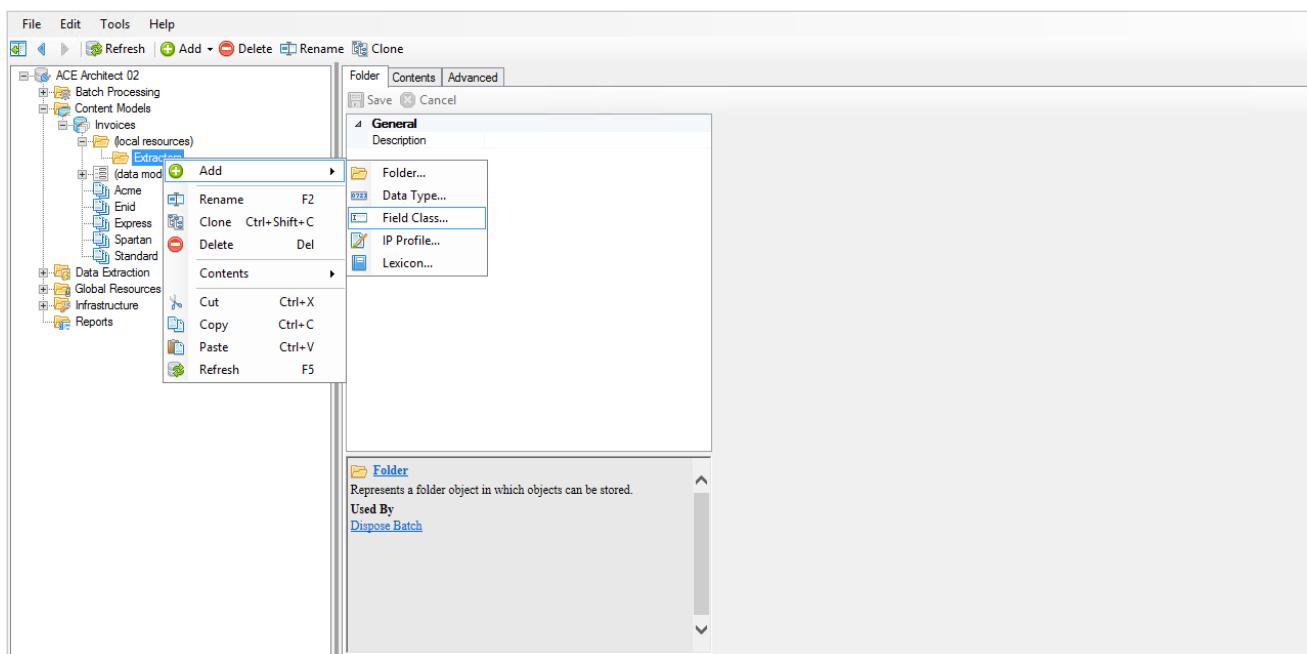




> Content Models / Invoices / (local resources) / Extractors

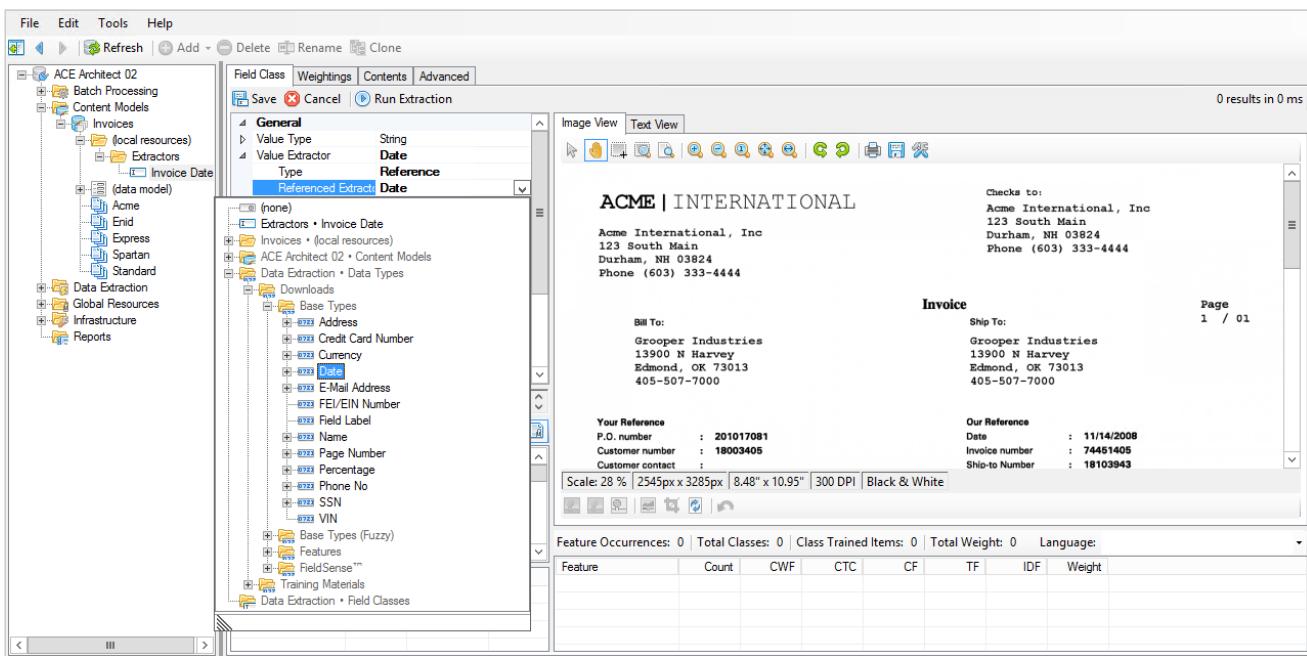
1. Right click and select Add > Field Class...

2. Name it **Invoice Date**



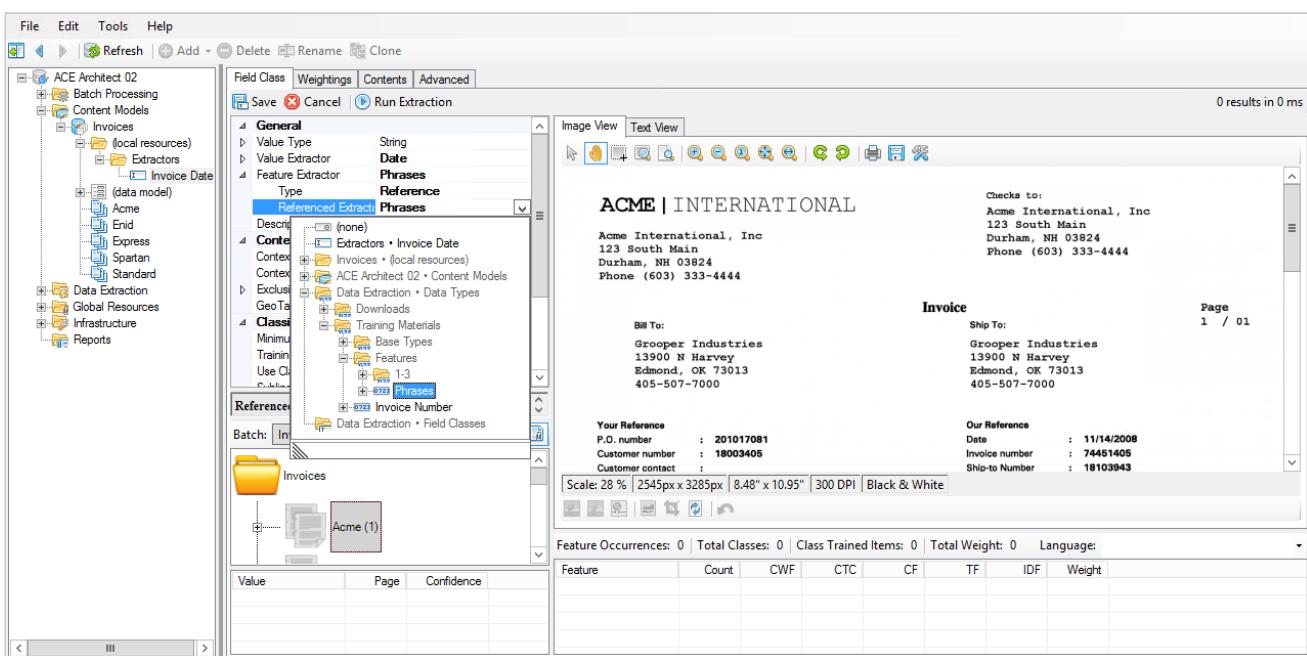
> Content Models / Invoices / (local resources) / Extractors / **Invoice Date** (Field Class)

1. In the General section, set the Value Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Data Extraction • Data Types > Downloads > Base Types > Date Data Type



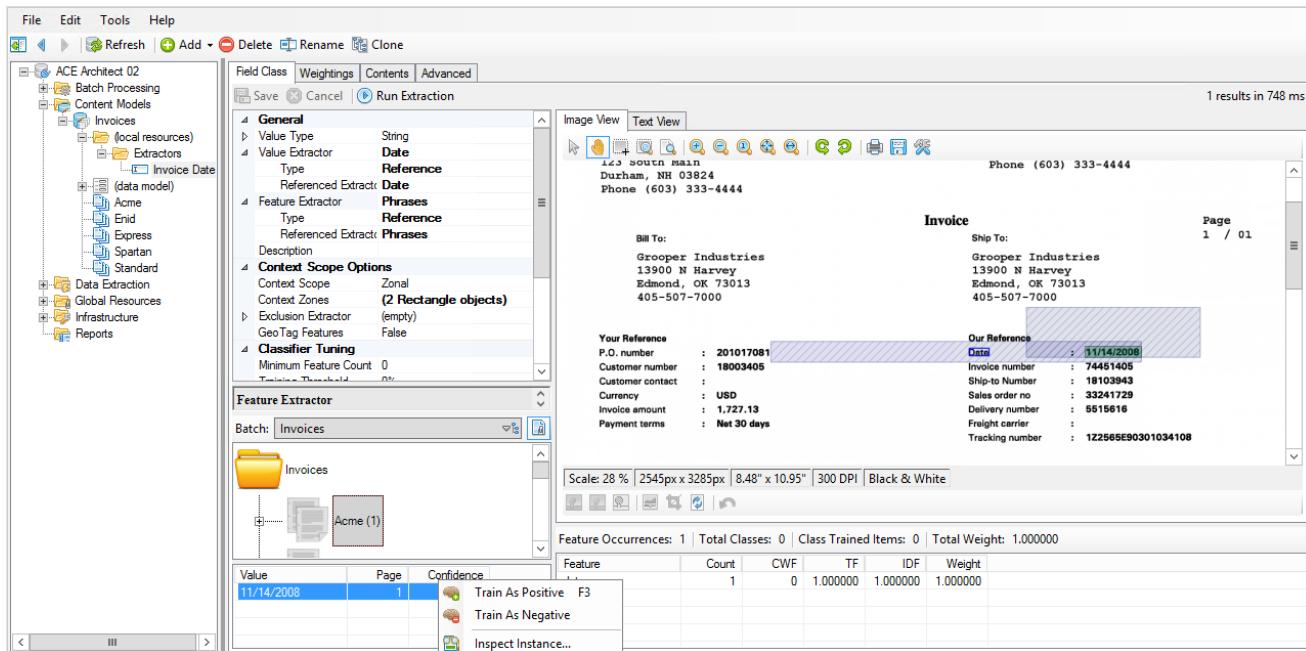
> Content Models / Invoices / (local resources) / Extractors / **Invoice Date**

1. In the General section, set the Feature Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Data Extraction • Data Types > Training Materials > Features > Phrases Data Type



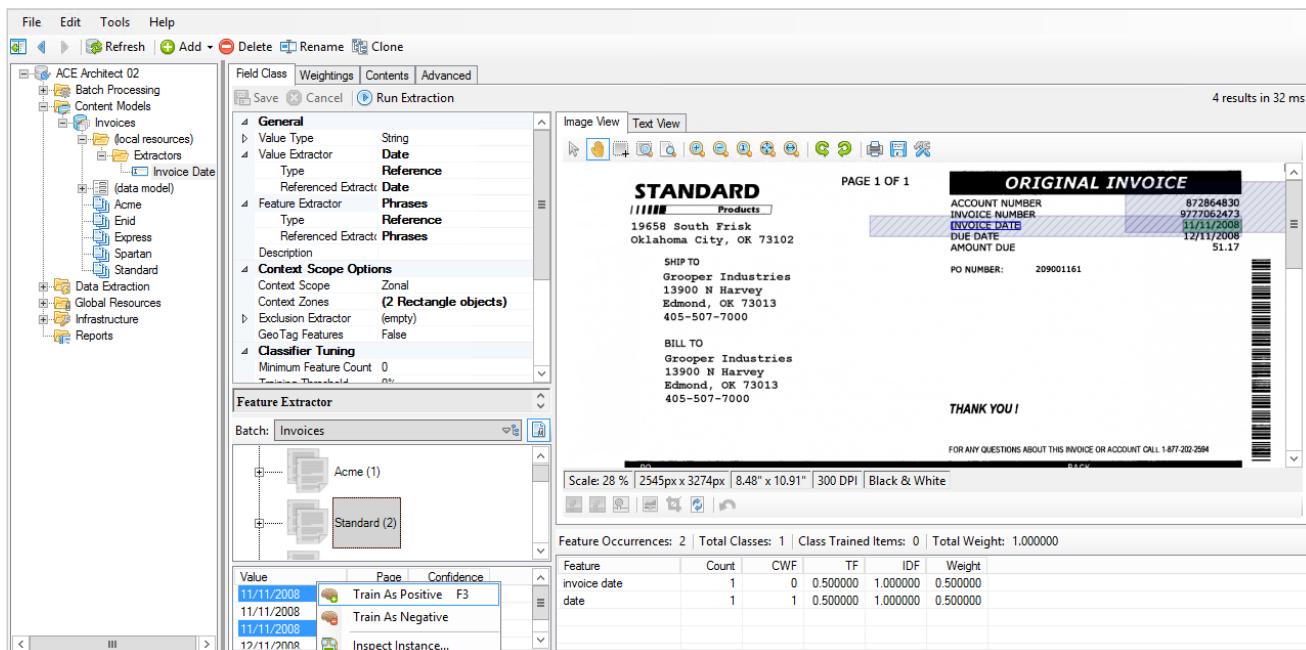
> Content Models / Invoices / (local resources) / Extractors / **Invoice Date**

1. Select Acme (1)
2. Click the Run Extraction button
3. In the candidate list, right click on the **11/14/2008** value and select Train As Positive



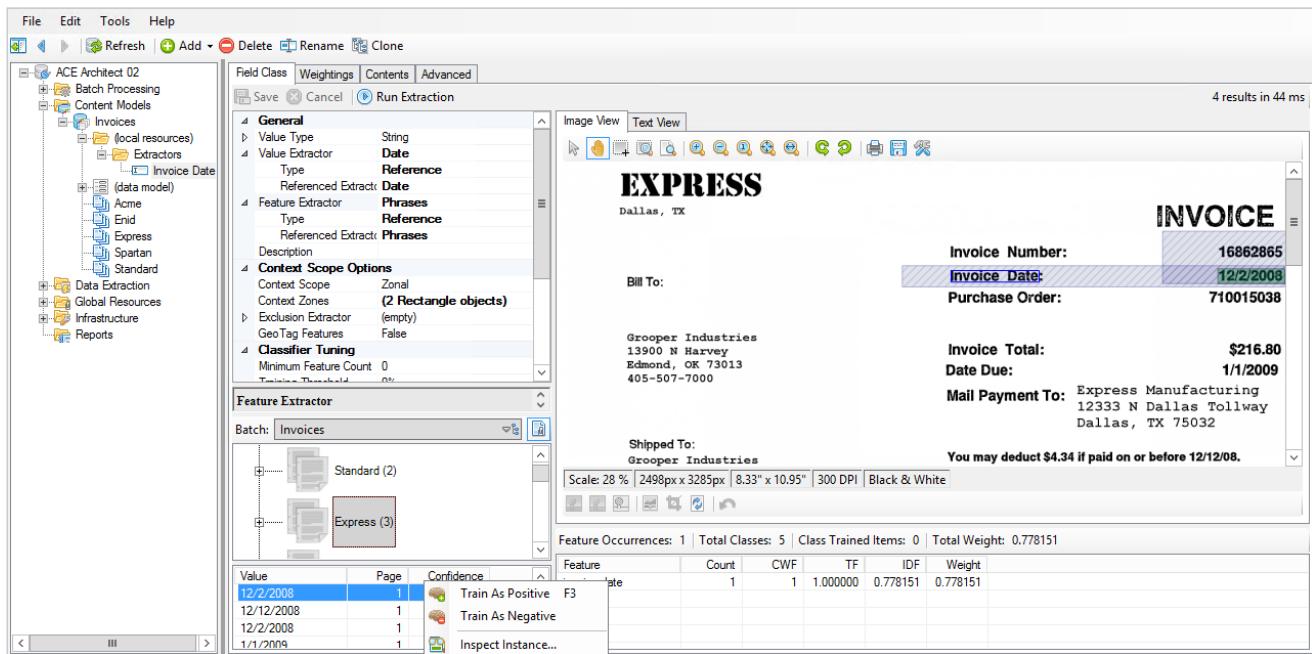
> Content Models / Invoices / (local resources) / Extractors / **Invoice Date**

1. Select Standard (2)
2. Select both of the **11/11/2008** candidates that have correct features
3. Right click on either of them and select Train As Positive



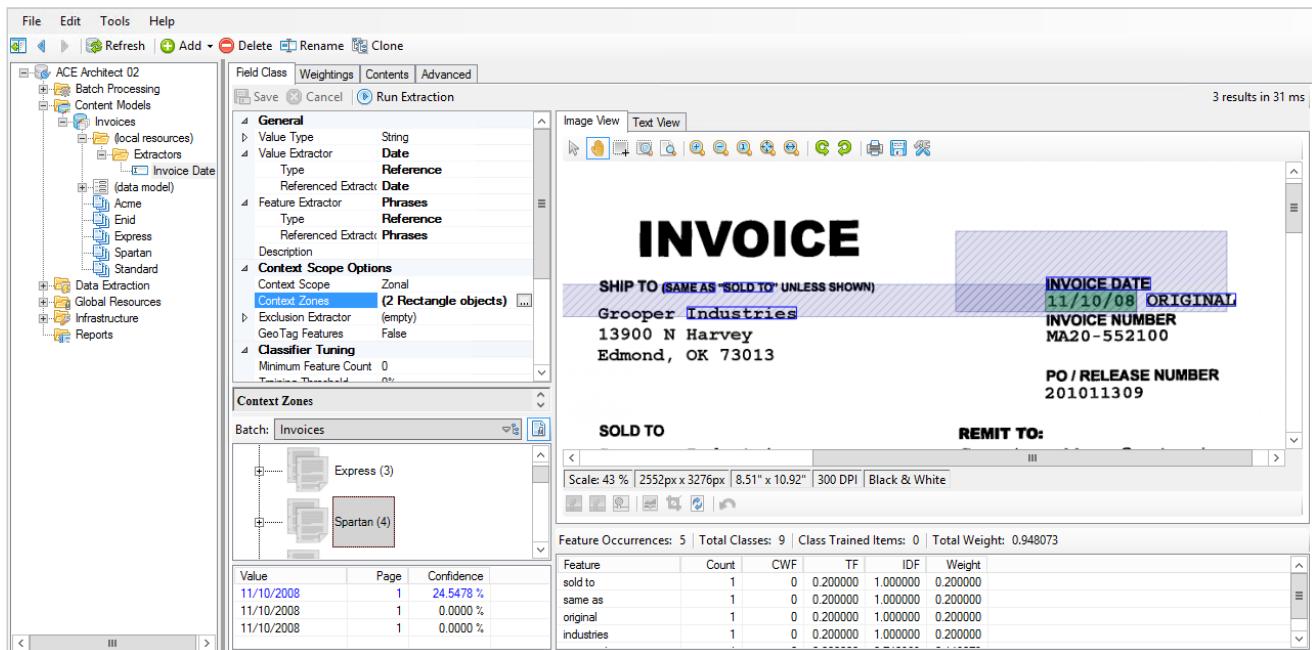
> Step

1. Select Express (3)
2. Right click on the **12/2/2008** candidate with 58.1689% confidence and select Train As Positive



> Step

1. Select Spartan (4)
2. Click to edit the Context Zones property



> Step

1. In the Context Zones window, edit the dimensions as follows:

a. Zone 1:

i. Left: -3.278

ii. Top: -0.054

iii. Right: 0.312

iv. Bottom: 0.054

b. Zone 2:

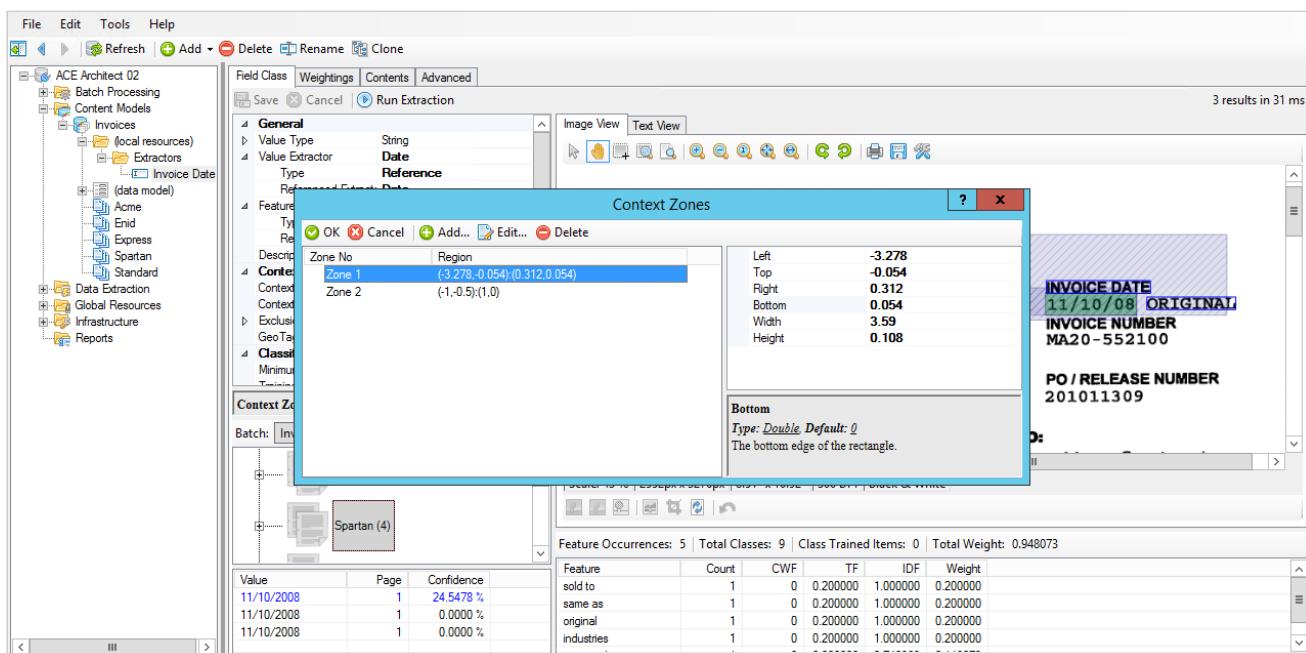
i. Left: -0.67

ii. Top: -0.37

iii. Right: 0.5

iv. Bottom: 0.04

2. Click OK



File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - Extractors
 - Invoice Date
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Field Class Weightings Contents Advanced

Save Cancel Run Extraction

3 results in 31 ms

Context Zones

OK Cancel Add... Edit... Delete

Zone No	Region	Left	-0.67
Zone 1	(-3.278,-0.054)(0.312,0.054)	Top	-0.37
Zone 2	(-0.67,-0.37)(0.5,0.04)	Right	0.5
		Bottom	0.04
		Width	1.17
		Height	0.41

Left
Type: Double Default: 0
The left edge of the rectangle.

Feature Occurrences: 5 | Total Classes: 9 | Class Trained Items: 0 | Total Weight: 0.948073

Feature	Count	CWF	TF	IDF	Weight
sold to	1	0	0.200000	1.000000	0.200000
same as	1	0	0.200000	1.000000	0.200000
original	1	0	0.200000	1.000000	0.200000
industries	1	0	0.200000	1.000000	0.200000

Value Page Confidence

11/10/2008	1	24.5478 %
11/10/2008	1	0.0000 %
11/10/2008	1	0.0000 %

Step

- Select Enid (5)
- In the Output section, set the Minimum Confidence property to 70%

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - Extractors
 - Invoice Date
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Field Class Weightings Contents Advanced

Save Cancel Run Extraction

3 results in 15 ms

Context Scope Options

Context Scope Zonal
(2 Rectangle objects)

Exclusion Extractor (empty)

GeoTag Features False

Classifier Tuning

Minimum Feature Count 0
Training Threshold 0%
Use Class Frequency False
Sublinear TF Scaling False
Smooth IDF True

Output

Minimum Confidence 0%
Collation Method Individual
Order By Confidence
Direction Descending

Minimum Confidence

Batch: Invoices

Spartan (4)	Enid (5)
-------------	----------

Value Page Confidence

12/5/2008	1	70.7107 %
11/11/2008	1	42.0751 %
12/3/2008	1	42.0751 %

INVOICE

INVOICE DATE: 12-05-08 INVOICE #: CSI/10076138 PAGE: 1

PH: (405) 555-1212 FAX: (405) 444-2121

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

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

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

ORDER NUMBER	ORDER DATE	CUSTOMER NUMBER	SALES PERSON	PURCHASE ORDER NUMBER	SHIP VIA	SHIP DATE	TERMS
100000755	11-11-08	100001410	000275	0100012704	FWD	12-03-08	Net 30

Scale: 28 % | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

Feature Occurrences: 1 | Total Classes: 9 | Class Trained Items: 0 | Total Weight: 0.740363

Feature	Count	CWF	TF	IDF	Weight
invoice date	1	2	1.000000	0.740363	0.740363

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - Extractors
 - Invoice Date
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Field Class Weightings Contents Advanced

Save Cancel Run Extraction

1 results in 17 ms

Context Scope Options

Context Scope Zonal
(2 Rectangle objects)

Exclusion Extractor (empty)

GeoTag Features False

Classifier Tuning

Minimum Feature Count 0
Training Threshold 0%
Use Class Frequency False
Sublinear TF Scaling False
Smooth IDF True

Output

Minimum Confidence 70%
Collation Method Individual
Order By Confidence
Direction Descending

Minimum Confidence

Batch: Invoices

Spartan (4)	Enid (5)
-------------	----------

Value Page Confidence

12/5/2008	1	70.7107 %
-----------	---	-----------

INVOICE

INVOICE DATE: 12-05-08 INVOICE #: CSI/10076138 PAGE: 1

PH: (405) 555-1212 FAX: (405) 444-2121

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

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

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

ORDER NUMBER	ORDER DATE	CUSTOMER NUMBER	SALES PERSON	PURCHASE ORDER NUMBER	SHIP VIA	SHIP DATE	TERMS
100000755	11-11-08	100001410	000275	0100012704	FWD	12-03-08	Net 30

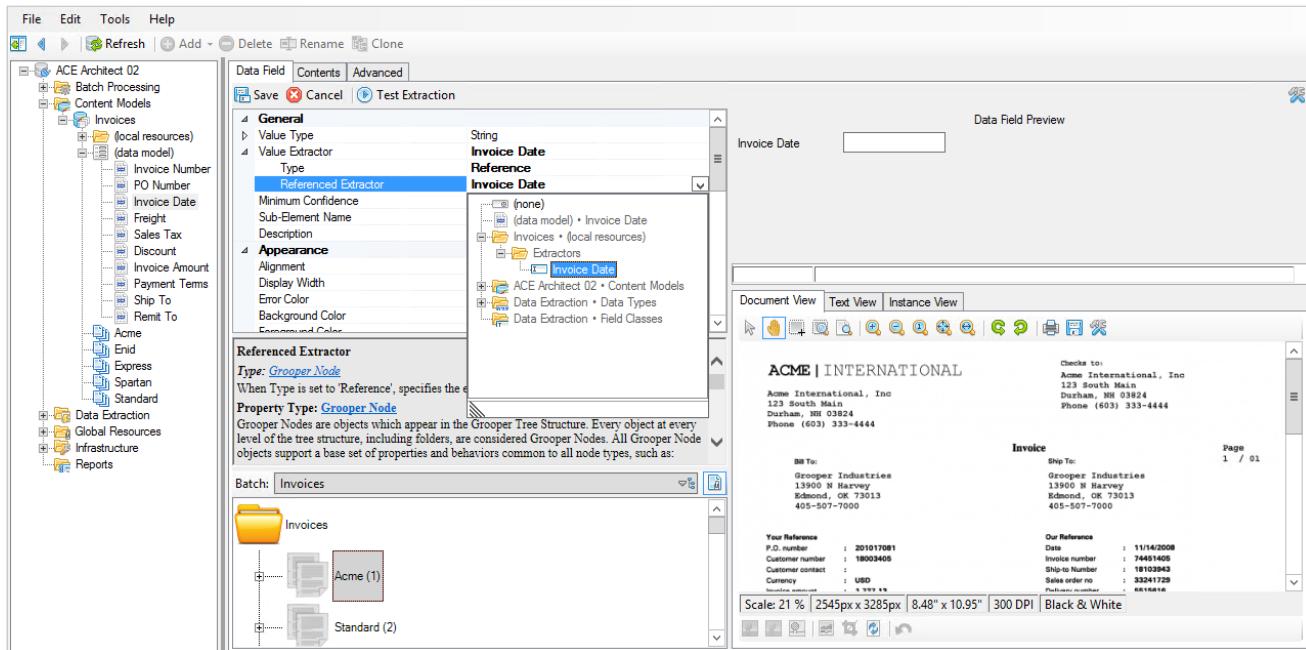
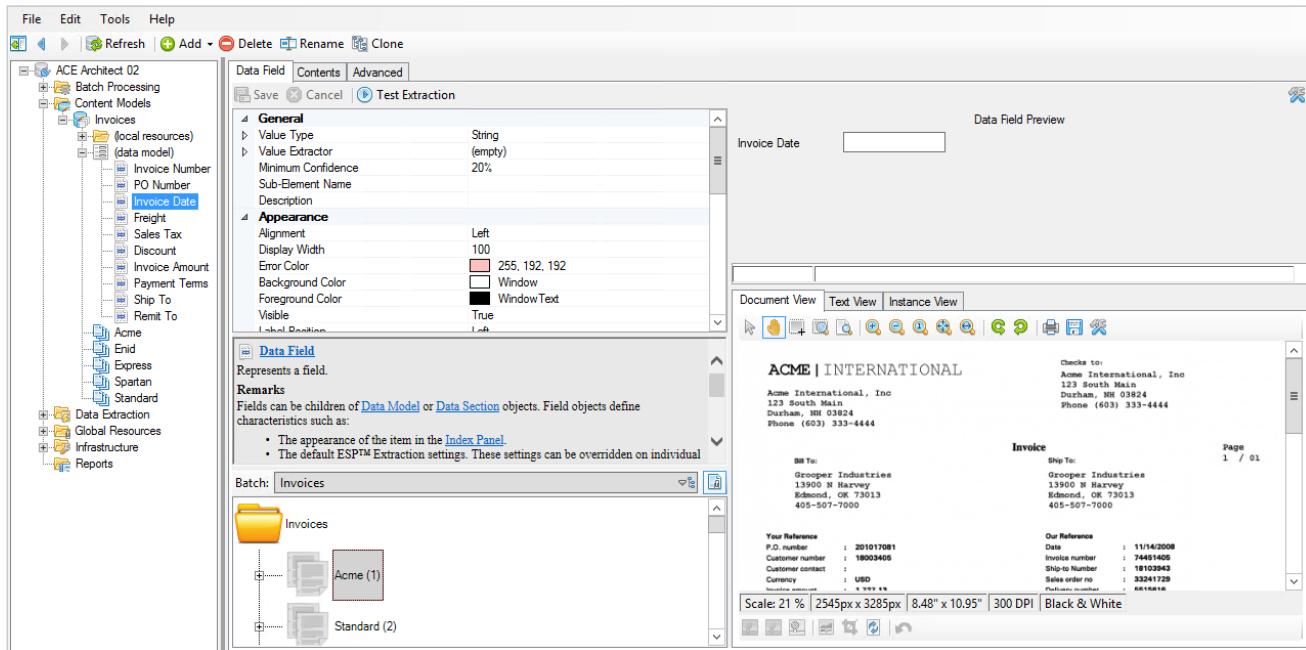
Scale: 28 % | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

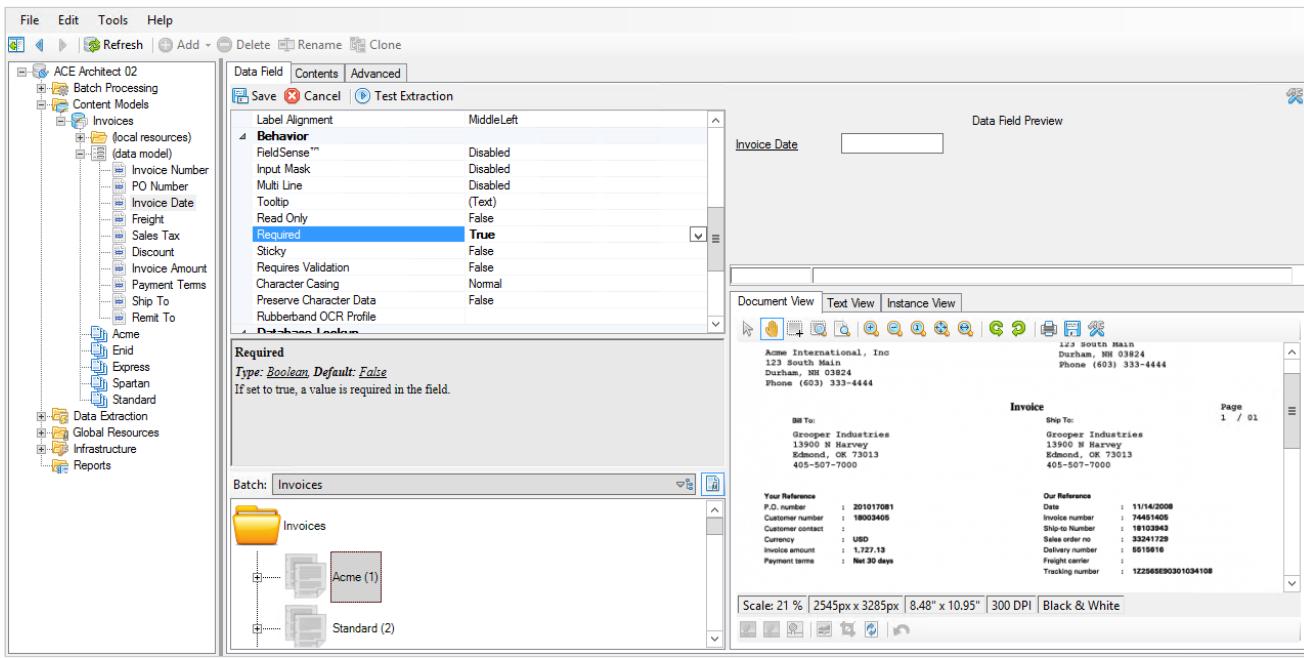
Feature Occurrences: 1 | Total Classes: 9 | Class Trained Items: 0 | Total Weight: 0.740363

Feature	Count	CWF	TF	IDF	Weight
invoice date	1	2	1.000000	0.740363	0.740363

Step

1. In the Data Model, select the Invoice Date Data Field
2. In the General section, set the Value Extractor > Type property to Reference
3. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > Invoice Date Field Class.
4. In the Behavior section, set the Required property to **True**

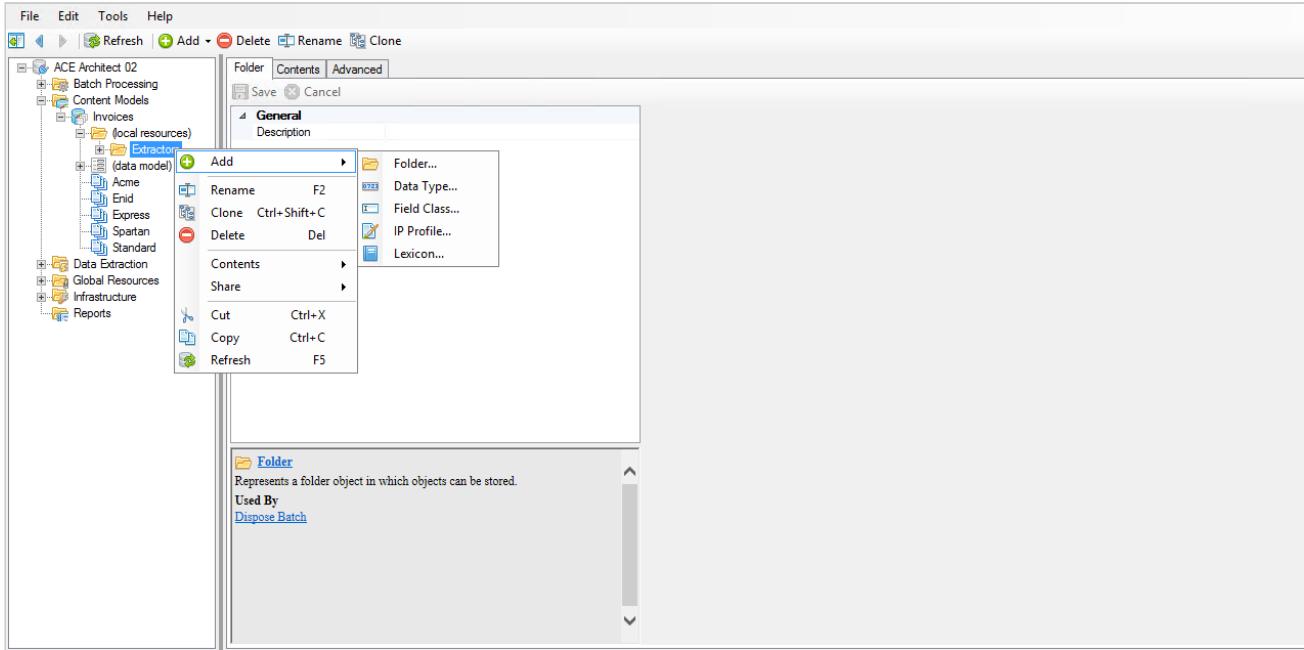




Ship To

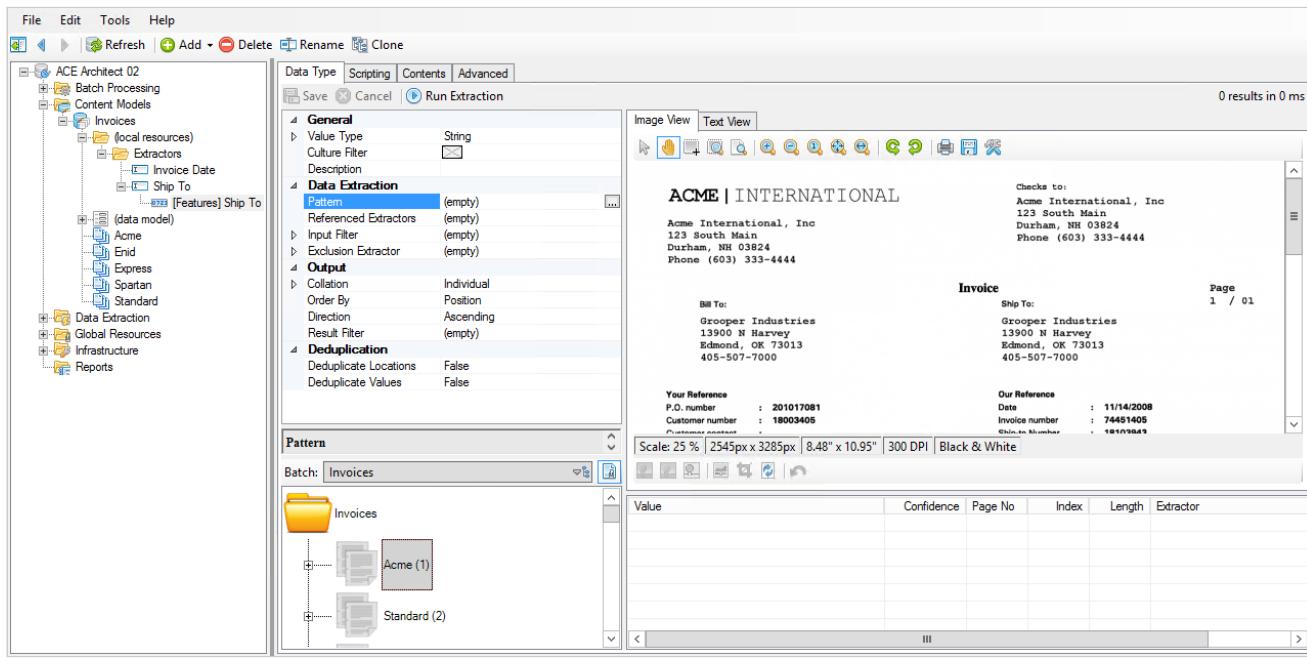
Step

1. Right click on the Extractors folder and select Add > Field Class...
2. Name it **Ship To**



Step

1. Right click on the Ship To Field Class and select Add > Data Type...
2. Name it [Features] Ship To
3. In the Data Extraction section, click to edit the Pattern



Step

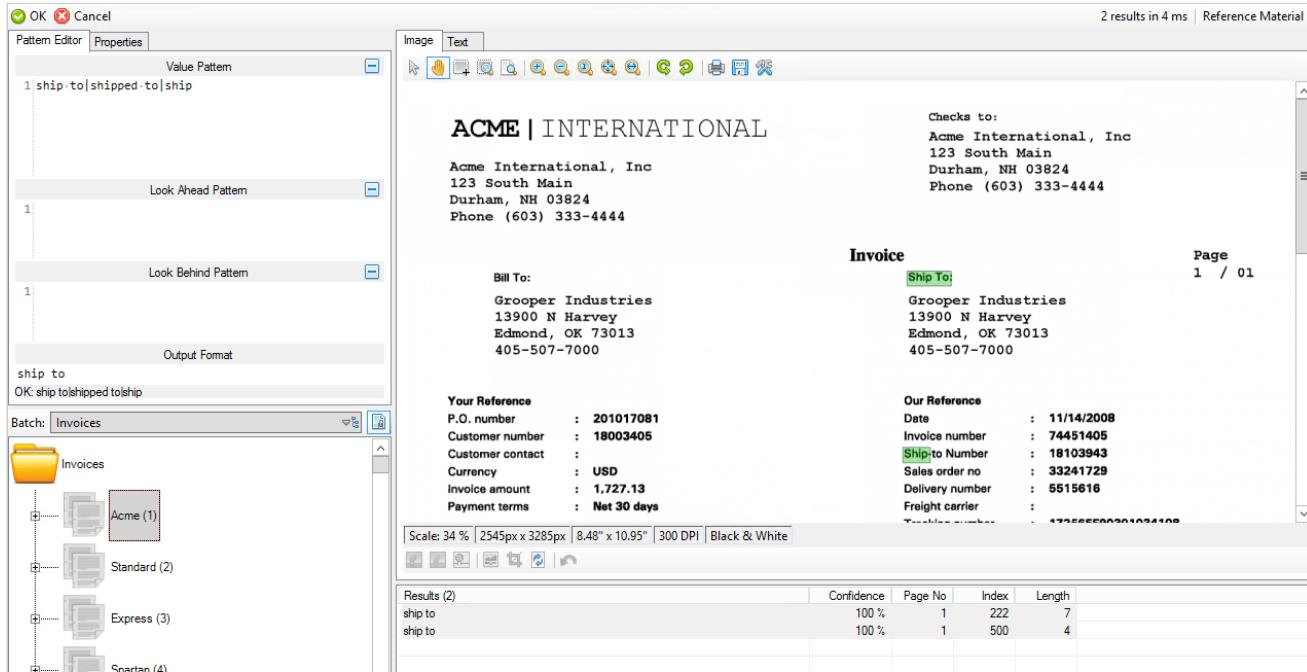
1. In the Value Pattern editor, type:

```
ship to|shipped to|ship
```

2. In the Output Format editor, type:

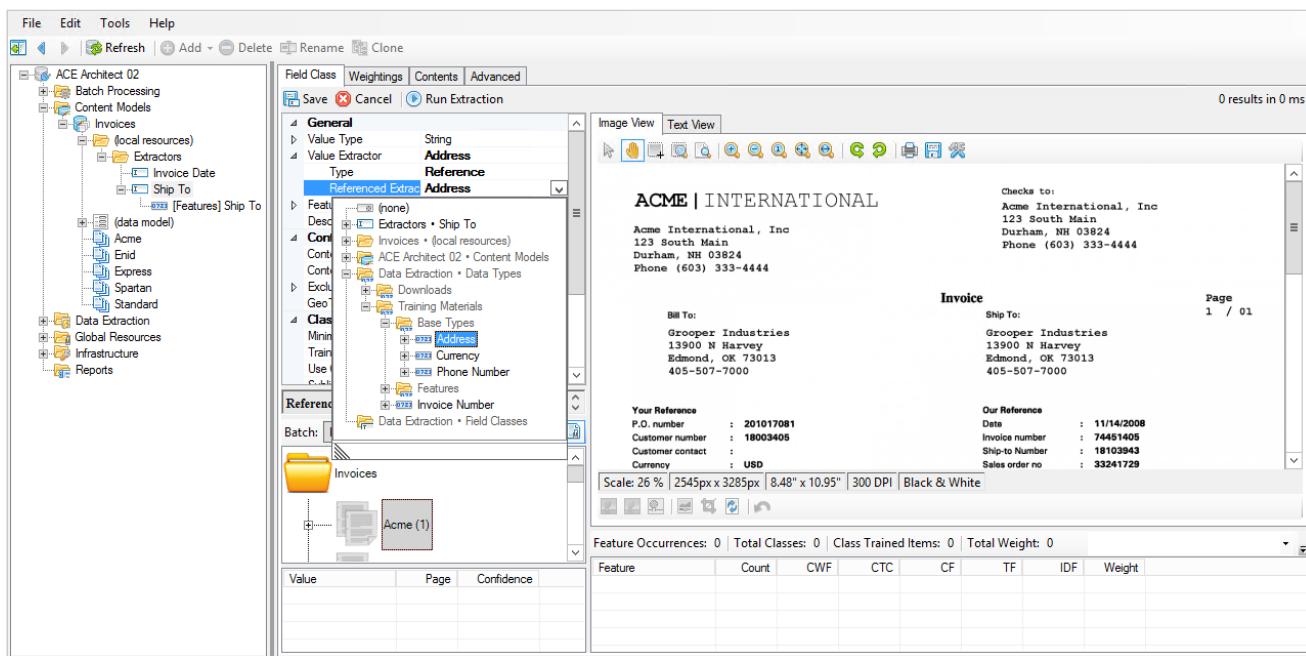
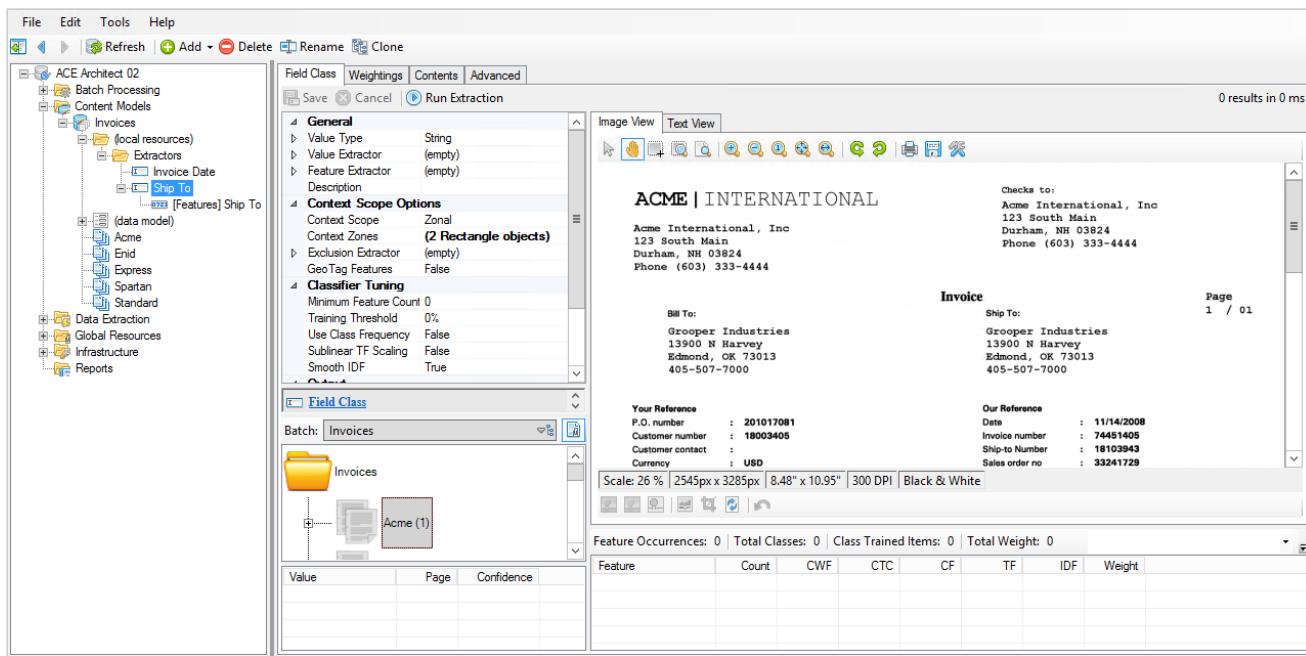
```
ship to
```

3. Click OK



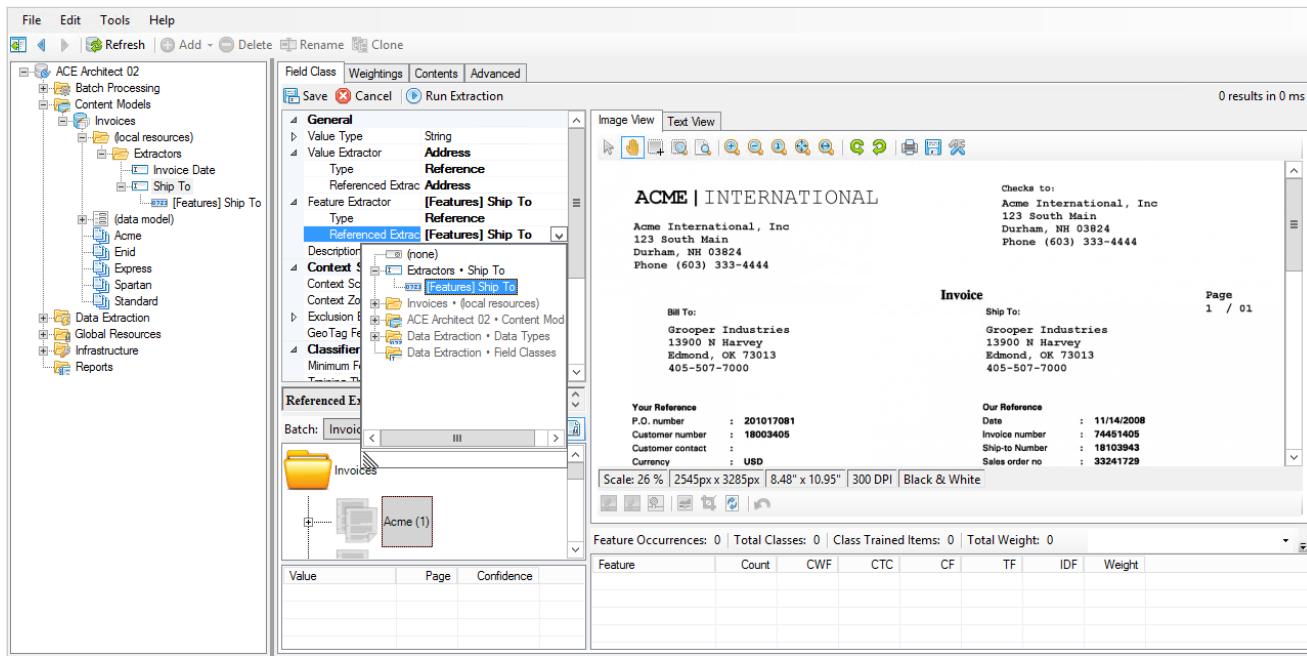
Step

1. Click on the Ship To Field Class
2. In the General section, set the Value Extractor > Type property to Reference
3. Set the Referenced Extractor property to the Data Extraction • Data Types > Training Materials > Base Types > Address Data Type



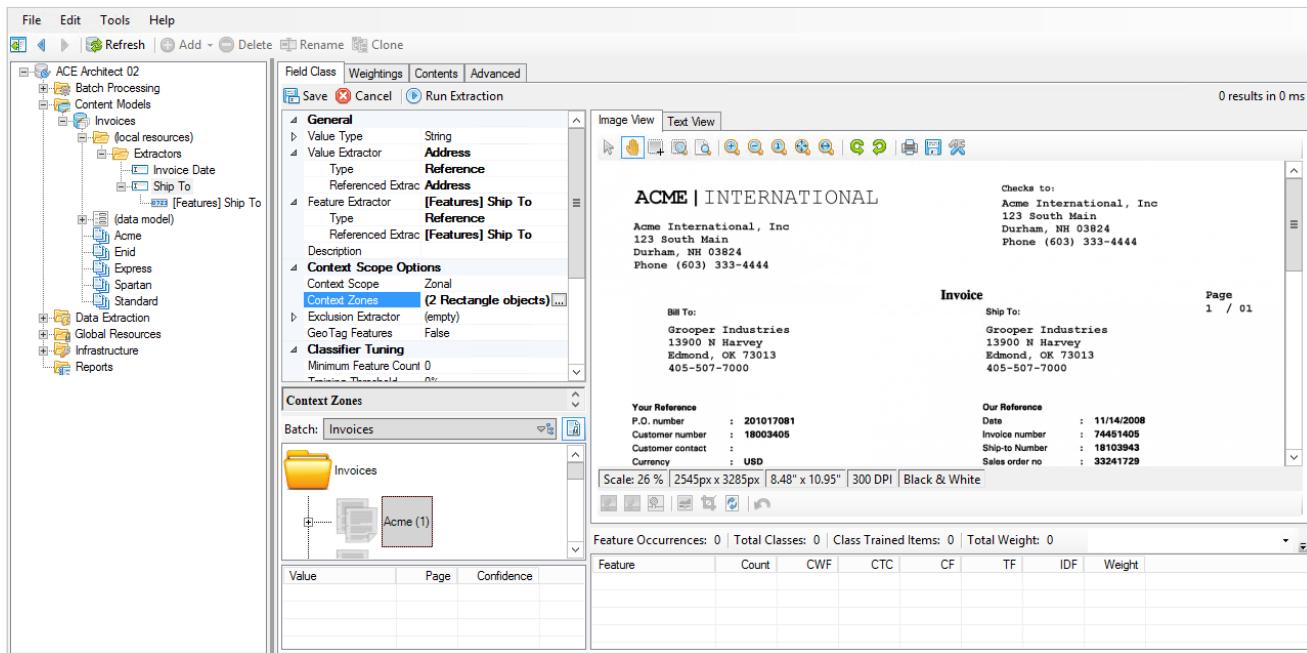
Step

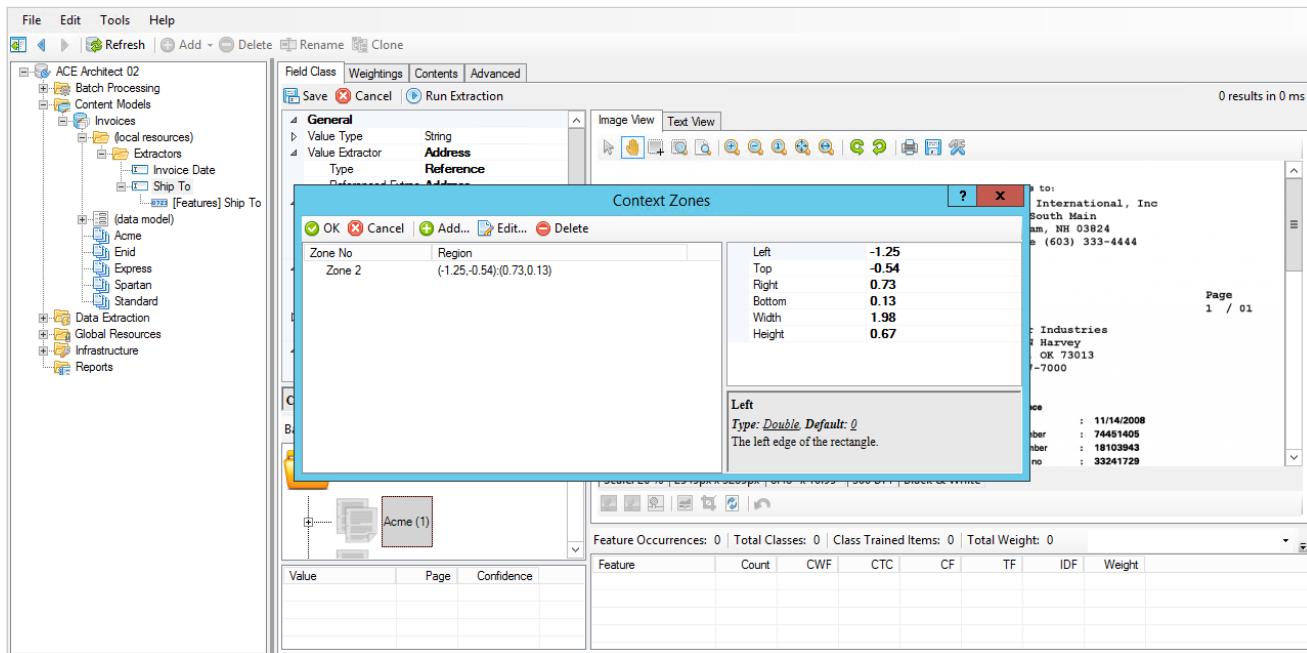
1. In the General section, set the Value Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Extractors • Ship To > [Features] Ship To Data Type



Step

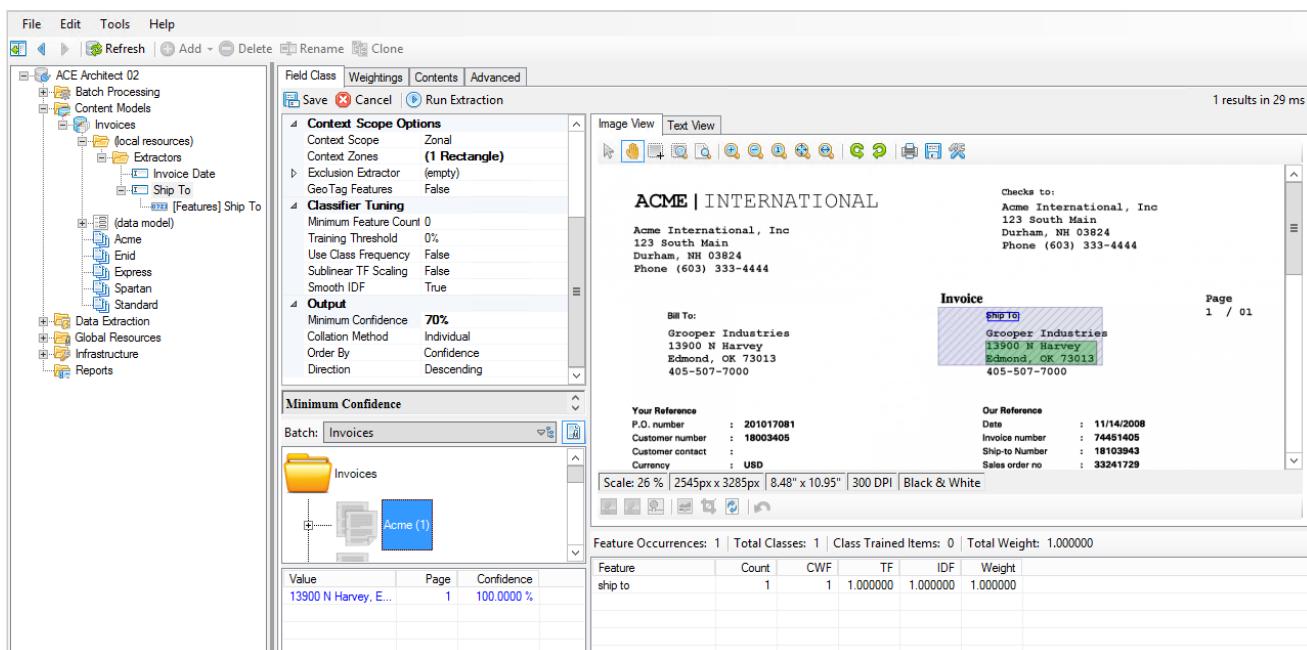
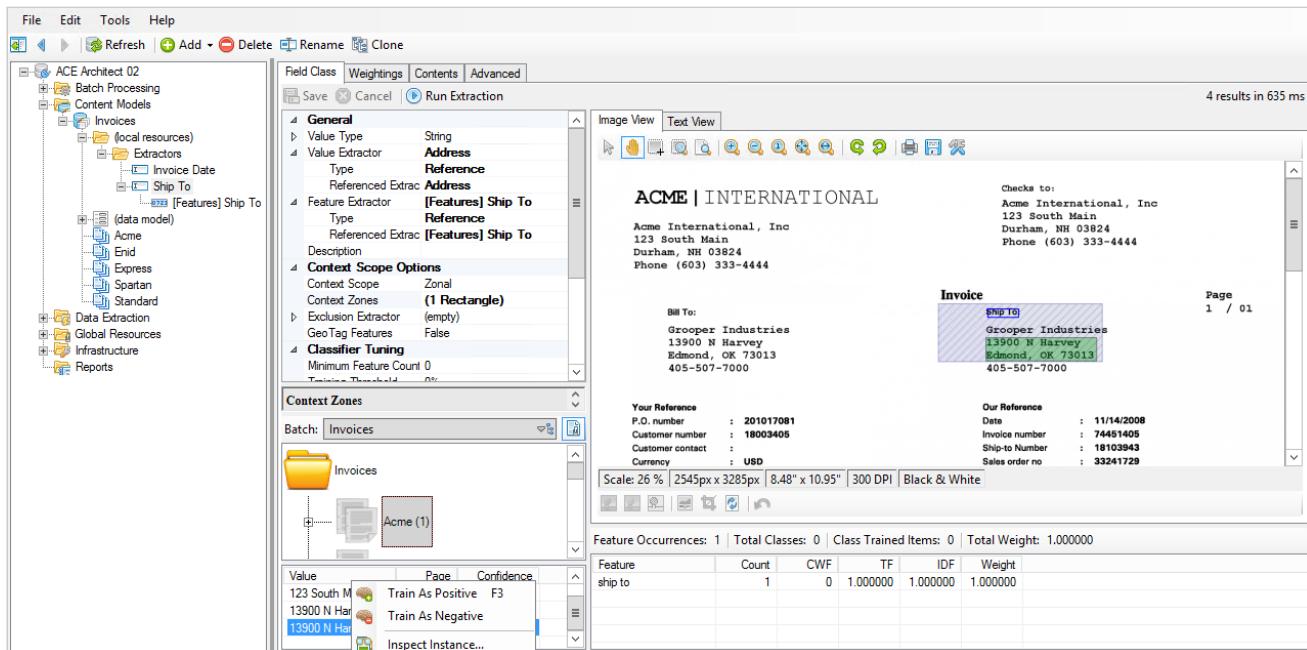
1. In the Context Scope Options section, click to edit the Context Zones
2. In the Context Zones window, delete one of the zones
3. Edit the dimensions of the remaining zone:
 - a. Left: **-1.25**
 - b. Top: **-0.54**
 - c. Right: **0.73**
 - d. Bottom: **0.13**
4. Click OK
5. Click Run Extraction





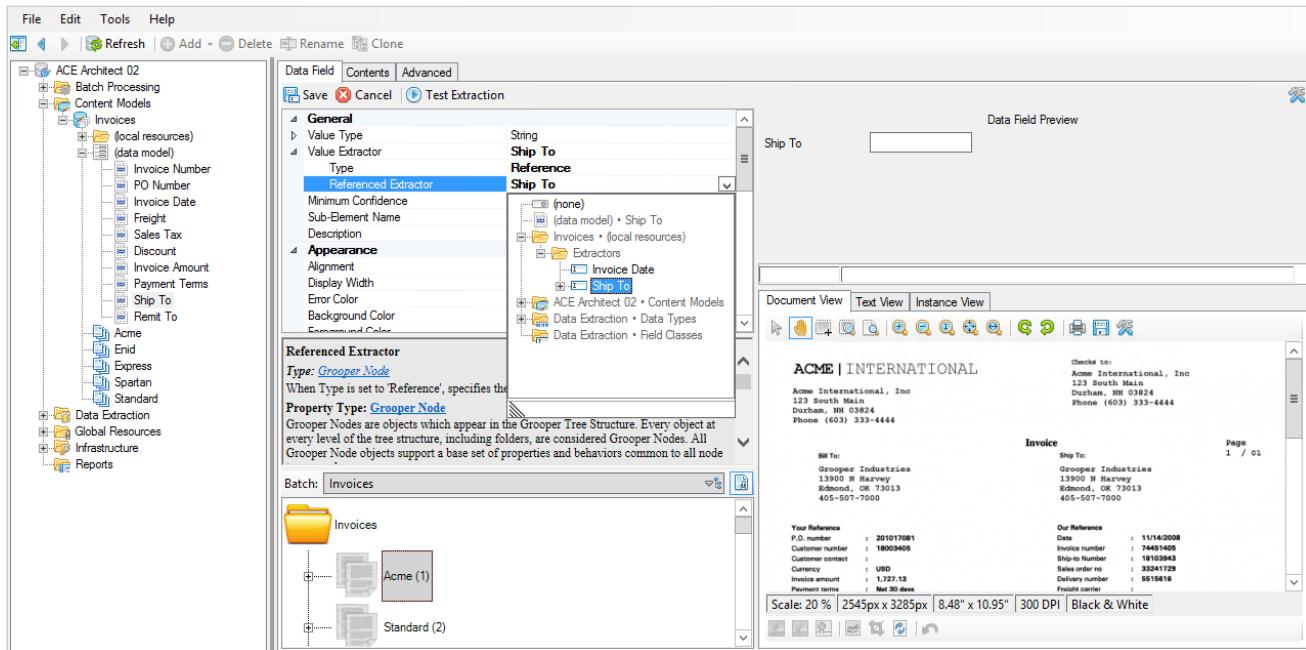
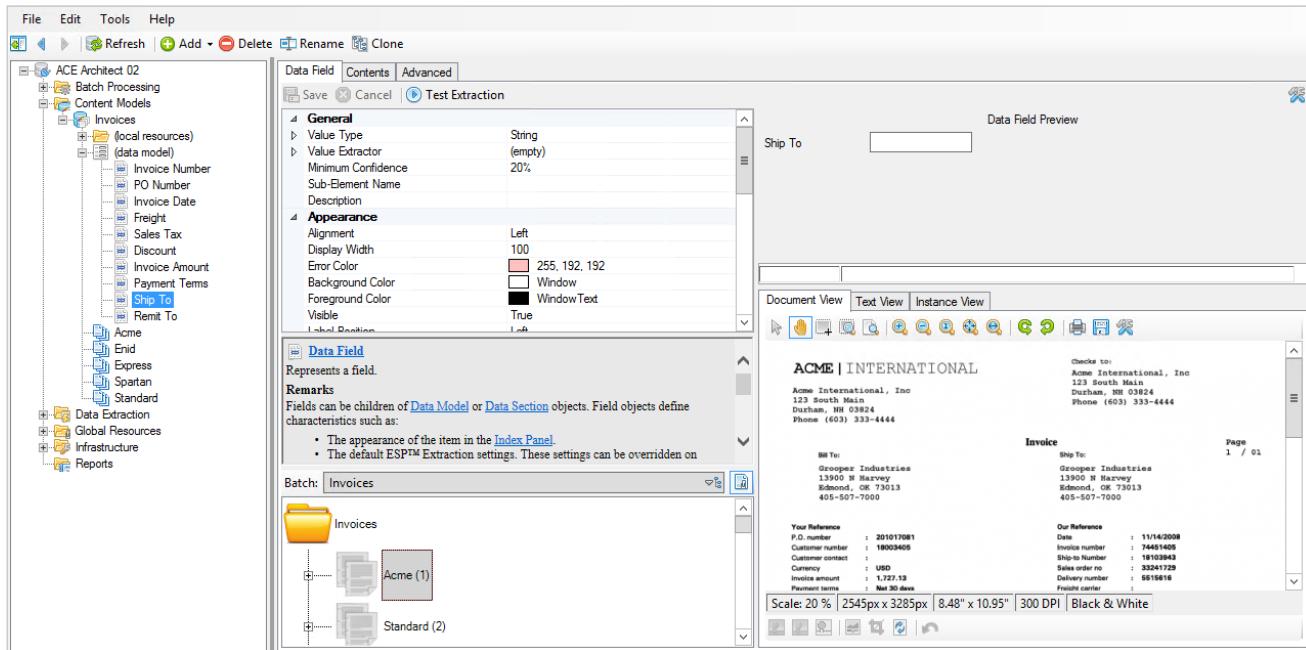
Step

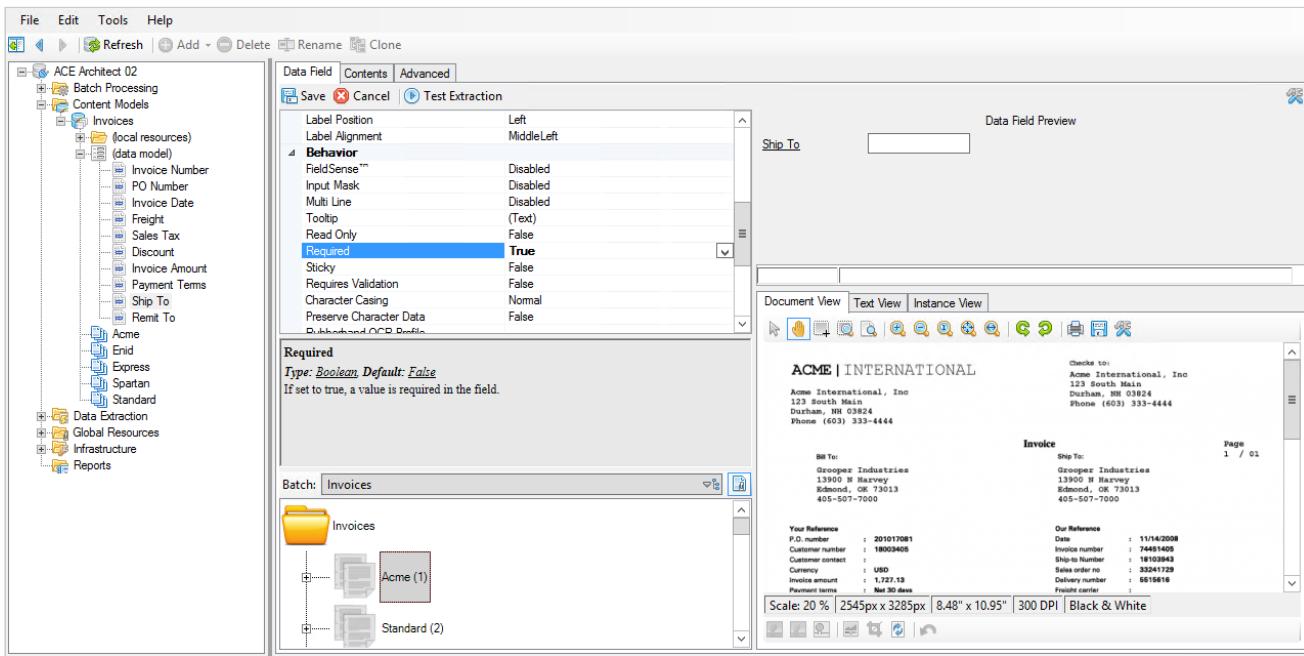
1. In the candidate list, find the value that has features that can be trained
2. Right click on that value and select Train As Positive
3. Set the Minimum Confidence property to **70%**



Step

1. In the Data Model, select the Ship To Data Field
2. In the General section, set the Value Extractor > Type property to Reference
3. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > Ship To Field Class.
4. In the Behavior section, set the Required property to **True**

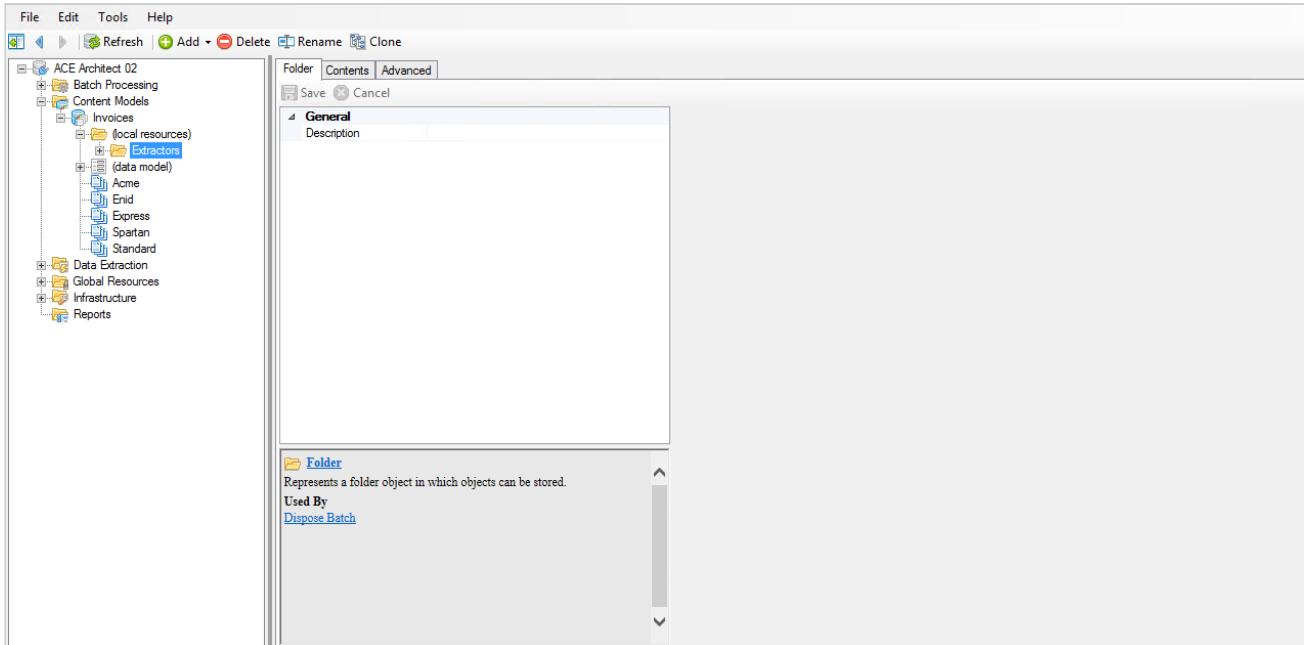


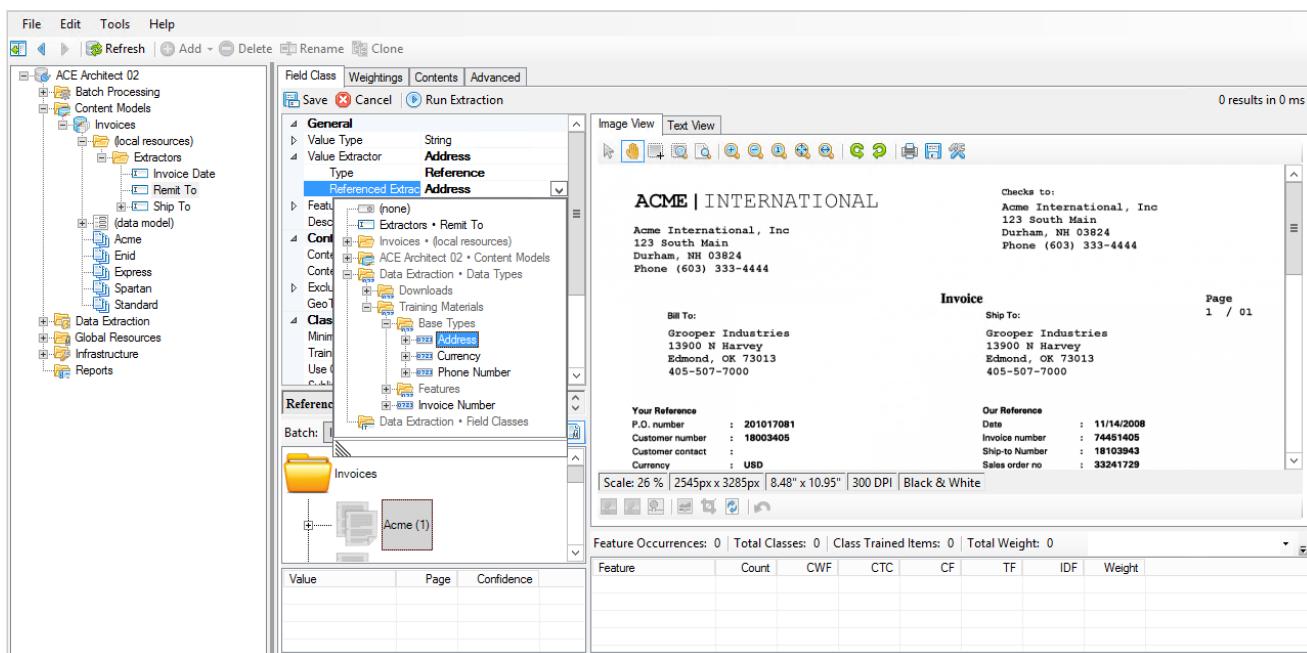
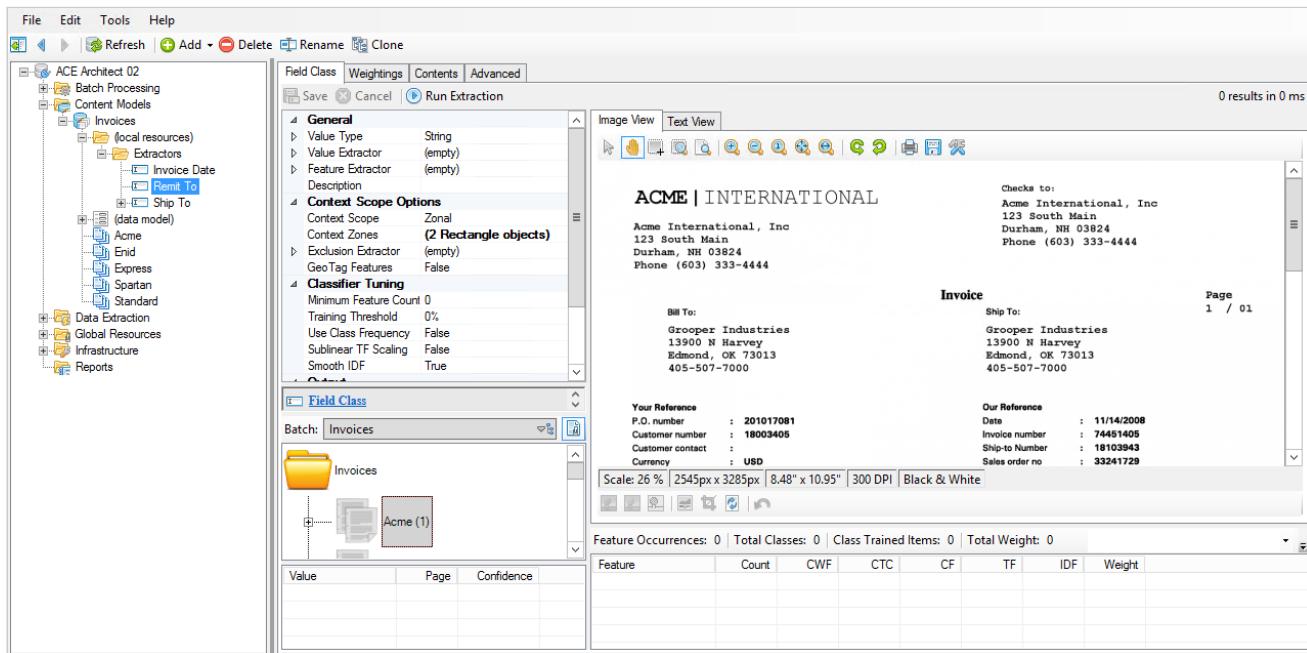


Remit To

Step

1. Right click on the Extractors folder and select Add > Field Class...
2. Name it Remit To
3. In the General section, set the Value Extractor > Type property to Reference
4. Set the Referenced Extractor property to the Data Extraction • Data Types > Training Materials > Base Types > Address Data Type

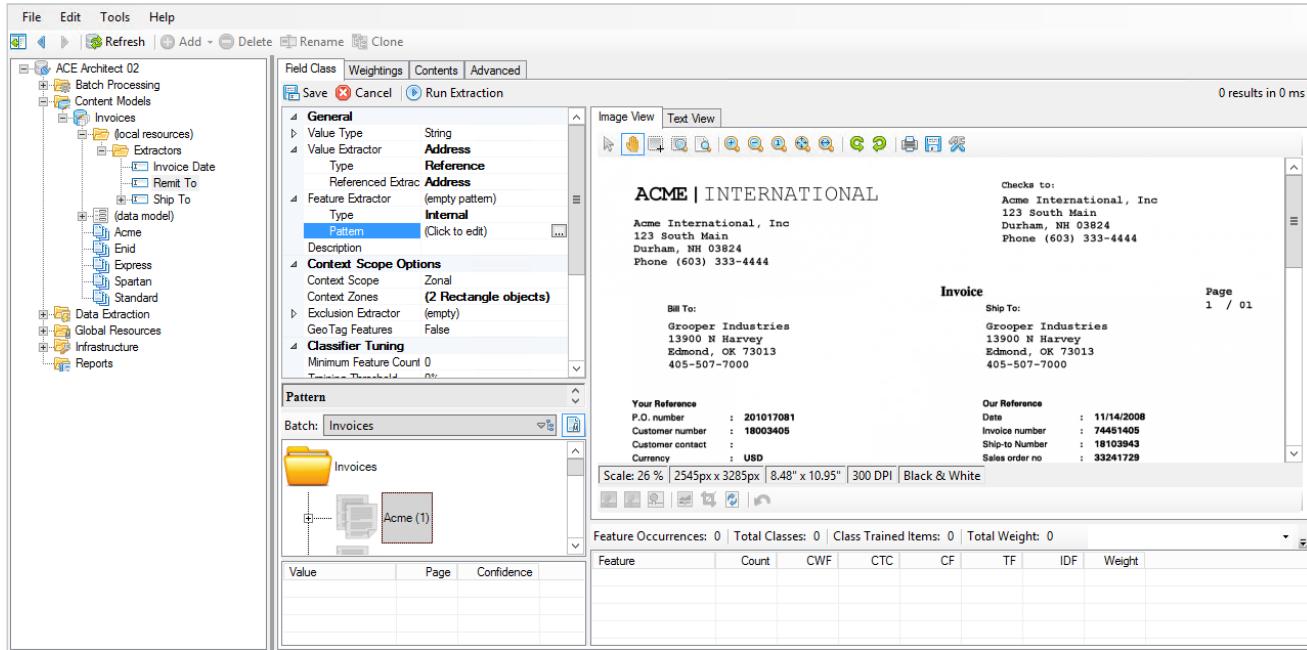




Step

1. In the General section, set the Feature Extractor > Type property to Internal

2. Click to edit the Pattern



Step

1. In the Value Pattern editor, type:

```
remit to|mail payment to|checks to
```

2. In the Output Format editor, type:

```
remit to
```

3. Select Standard (2)

4. Click on the Properties tab

5. Change the Mode property to FuzzyRegEx

6. In the Fuzzy Matching Options section, expand Fuzzy Match Weightings

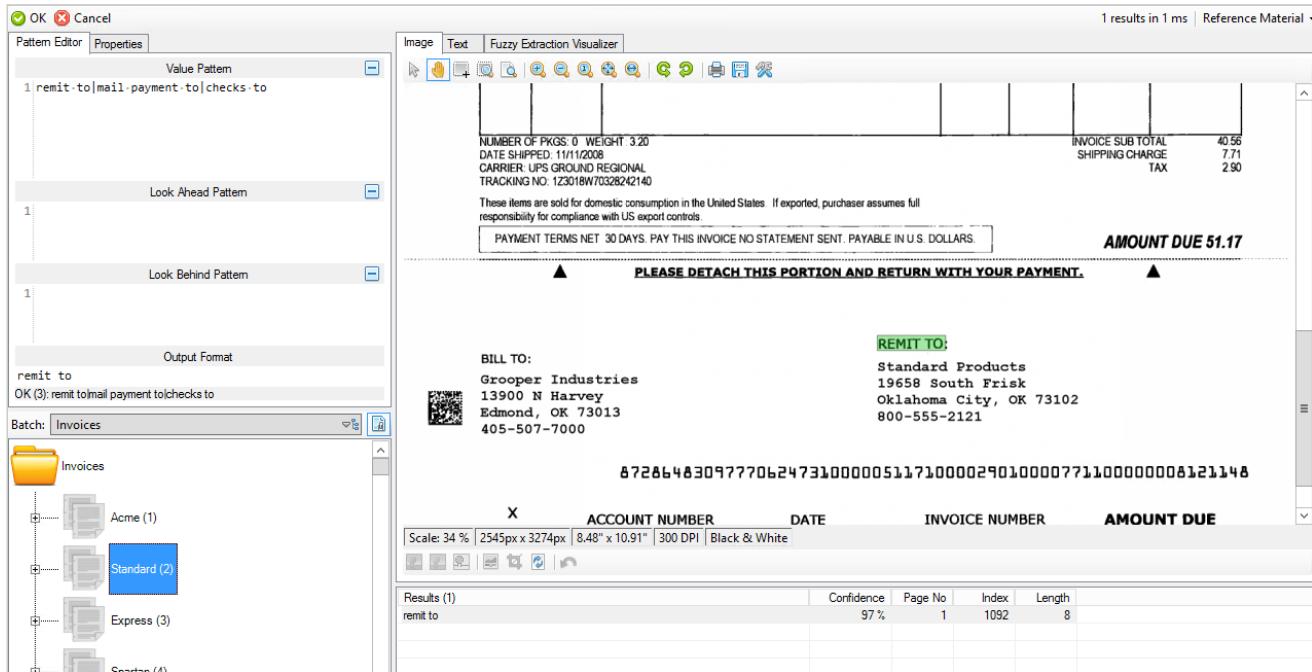
7. Click to edit Local Entries

8. In the List Editor window, type:

```
li=0.25
```

9. Click OK to close the List Editor

10. Click OK to close the Pattern Editor



Pattern Editor

Properties

General

- Value Type: String
- Mode: FuzzyRegEx
- Case Sensitive: False
- Preprocessing Options: None
- Regional Settings: (empty)
- Expression Lexicon: (empty)
- Referenced Lexicons: (empty)

Fuzzy Matching Options

- Minimum Similarity: 90%
- Match Mode: LeastCost
- Fuzzy Match Weightings: (empty)

Regular Expression

Mode

Type: ExtractionMode Default: RegEx

Specifies the extraction mode. Can be one of the following values:

- RegEx - Normal regular expression mode. Finds all instances which are a 100% match for the regular expression. RegEx mode supports the full syntax and feature set of Microsoft .Net Framework Regular Expressions.

Batch: Invoices

Invoices

- Acme (1)
- Standard (2)
- Express (3)
- Spartan (4)

Text

PAGE 1 OF 1 ORIGINAL INVOICE<\r><\n>
STANDARD<\r><\n>ACCOUNT NUMBER 872864830<\r><\n>111EEm Products<\r><\n>INVOICE NUMBER 9777062473<\r><\n>19658 South Frisk INVOICE DATE 11/11/2008<\r><\n>Oklahoma City, OK 73102 DUE DATE 12/11/2008<\r><\n>AMOUNT DUE 51.17<\r><\n>SHIP TO<\r><\n>PO NUMBER: 209001161<\r><\n>Grooper Industries<\r><\n>13900 N Harvey<\r><\n>Edmond, OK 73013<\r><\n>405-507-7000<\r><\n>THANK YOU!<\r><\n>FOR ANY QUESTIONS ABOUT THIS INVOICE OR ACCOUNT CALL 1-877-202-2594<\r><\n>26526 FLASHLIGHT,D 8 5.07 40.56<\r><\n>MANUFACTURER # IV2DM<\r><\n>NUMBER OF PKGS: 0 WE GHT: 3.20 INVOICE SUB TOTAL 40.56<\r><\n>DATE SHIPPED: 11/11/2008 SHIPPING CHARGE 7.71<\r><\n>CARRIER: UPS GROUND REGIONAL TAX 2.90<\r><\n>TRACKING NO: 123018W70328242140<\r><\n>These items are sold for domestic consumption in the United States If exported, purchaser assumes full<\r><\n>responsibility for compliance with US export controls.<\r><\n>PAYMENT TERMS NET 30 DAYS. PAY THIS INVOICE NO STATEMENT SENT. PAYABLE IN U.S. DOLLARS. AAFOUNTDUE 51.17<\r><\n>PLEASE DETACH THIS PORTION AND RETURN WITH YOUR PAYMENT. A<\r><\n>Characters: 1,495 | Lines: 45 | Current Line: 0 | Current Page: 1

Results (0)

Confidence	Page No	Index	Length

Pattern Editor

Properties

General

- Value Type: String
- Mode: FuzzyRegEx
- Case Sensitive: False
- Preprocessing Options: None
- Regional Settings: (empty)
- Expression Lexicon: (empty)
- Referenced Lexicons: (empty)

Fuzzy Matching Options

- Minimum Similarity: 85%
- Match Mode: LeastCost
- Fuzzy Match Weightings: (empty)
- Local Entries: 0 entries
- Included Lexicons: (0 Lexicon objects)

Regular Expression

Local Entries

Type: String

A list of local lexicon entries.

Remarks

The list should be formatted so that there is one entry per line. Use the = symbol to indicate replacement values.

Batch: Invoices

Invoices

- Acme (1)
- Standard (2)
- Express (3)
- Spartan (4)

Text

405-507-7000<\r><\n>BILL TO<\r><\n>Grooper Industries<\r><\n>13900 N Harvey<\r><\n>Edmond, OK 73013<\r><\n>405-507-7000<\r><\n>THANK YOU!<\r><\n>FOR ANY QUESTIONS ABOUT THIS INVOICE OR ACCOUNT CALL 1-877-202-2594<\r><\n>UNE # ITEM # DESCRIPTION -QUANTIT.Y . DRDRED UNIT PRICE TOTAL<\r><\n>26526 FLASHLIGHT,D 8 5.07 40.56<\r><\n>MANUFACTURER # IV2DM<\r><\n>NUMBER OF PKGS: 0 WE GHT: 3.20 INVOICE SUB TOTAL 40.56<\r><\n>DATE SHIPPED: 11/11/2008 SHIPPING CHARGE 7.71<\r><\n>CARRIER: UPS GROUND REGIONAL TAX 2.90<\r><\n>TRACKING NO: 123018W70328242140<\r><\n>These items are sold for domestic consumption in the United States If exported, purchaser assumes full<\r><\n>responsibility for compliance with US export controls.<\r><\n>PAYMENT TERMS NET 30 DAYS. PAY THIS INVOICE NO STATEMENT SENT. PAYABLE IN U.S. DOLLARS. AAFOUNTDUE 51.17<\r><\n>PLEASE DETACH THIS PORTION AND RETURN WITH YOUR PAYMENT. A<\r><\n>REMIT TO:<\r><\n>BILL TO:<\r><\n>Standard Products<\r><\n>Grooper Industries 19658 South Frisk<\r><\n>IR 13900 N Harvey Oklahoma City, OK 73102<\r><\n>Edmond, OK 73013 800-555-2121<\r><\n>405-507-7000<\r><\n>8728648309770624731000005117100000029010000771100000008121148<\r><\n>X<\r><\n>ACCOUNT NUMBER DATE INVOICE NUMBER AMOUNT DUE<\r><\n>872864830977062473 51.17<\r><\n>FOR COMMENTS OR CHANGE OF ADDRESS, ENTER INFORMATION ON REVERSE SIDE.<\r><\n>1 YNN 0030800 P

Results (1)

Confidence	Page No	Index	Length
88 %	1	1092	8

Pattern Editor

Properties

General

- Value Type: String
- Mode: FuzzyRegEx
- Case Sensitive: False
- Preprocessing Options: None
- Regional Settings: (empty)
- Expression Lexicon: (empty)
- Referenced Lexicons: (empty)

Fuzzy Matching Options

- Minimum Similarity: 85%
- Match Mode: LeastCost
- Fuzzy Match Weightings: (empty)
- Local Entries: 0 entries
- Included Lexicons: (0 Lexicon objects)

Regular Expression

Local Entries

Type: String

A list of local lexicon entries.

Remarks

The list should be formatted so that there is one entry per line. Use the = symbol to indicate replacement values.

Batch: Invoices

Invoices

- Acme (1)
- Standard (2)
- Express (3)
- Spartan (4)

Text

405-507-7000<\r><\n>BILL TO<\r><\n>List Editor

OK Cancel

1i=0.25|

<\n>

xported, purchaser assumes full<\r><\n>IN U.S. DOLLARS. AAFOUNTDUE 51.17<\r><\n>r><\n>

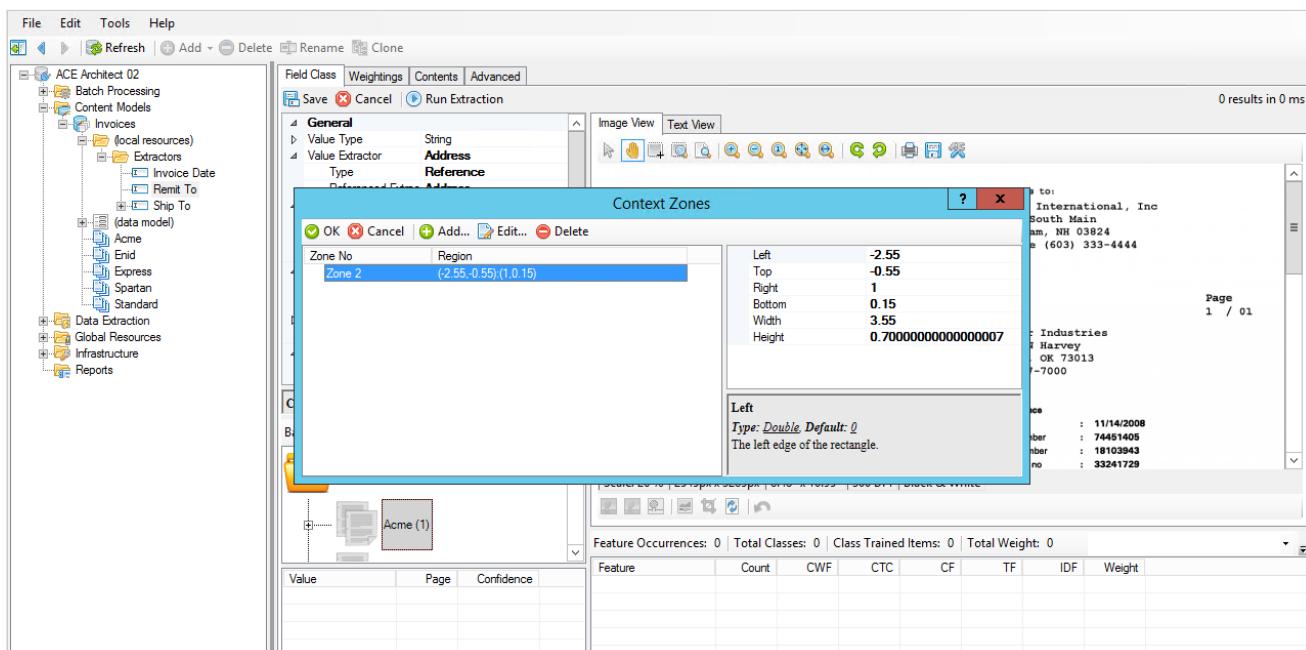
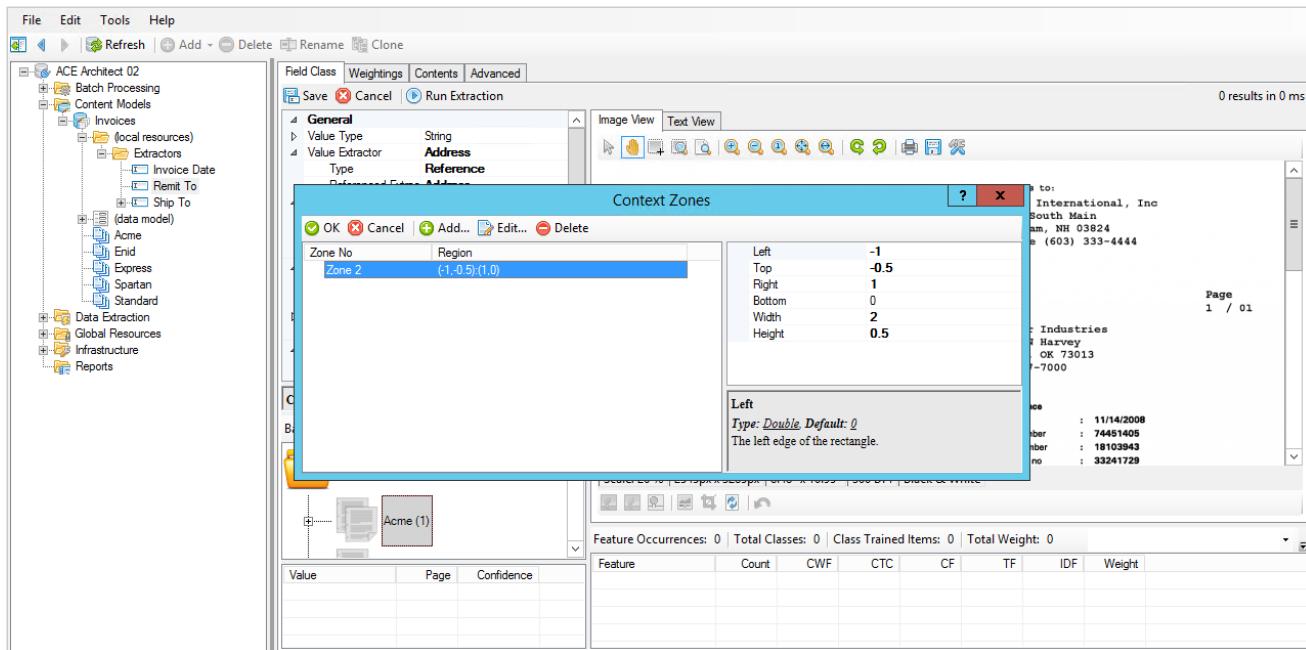
Results (1)

Confidence	Page No	Index	Length
88 %	1	1092	8

Step

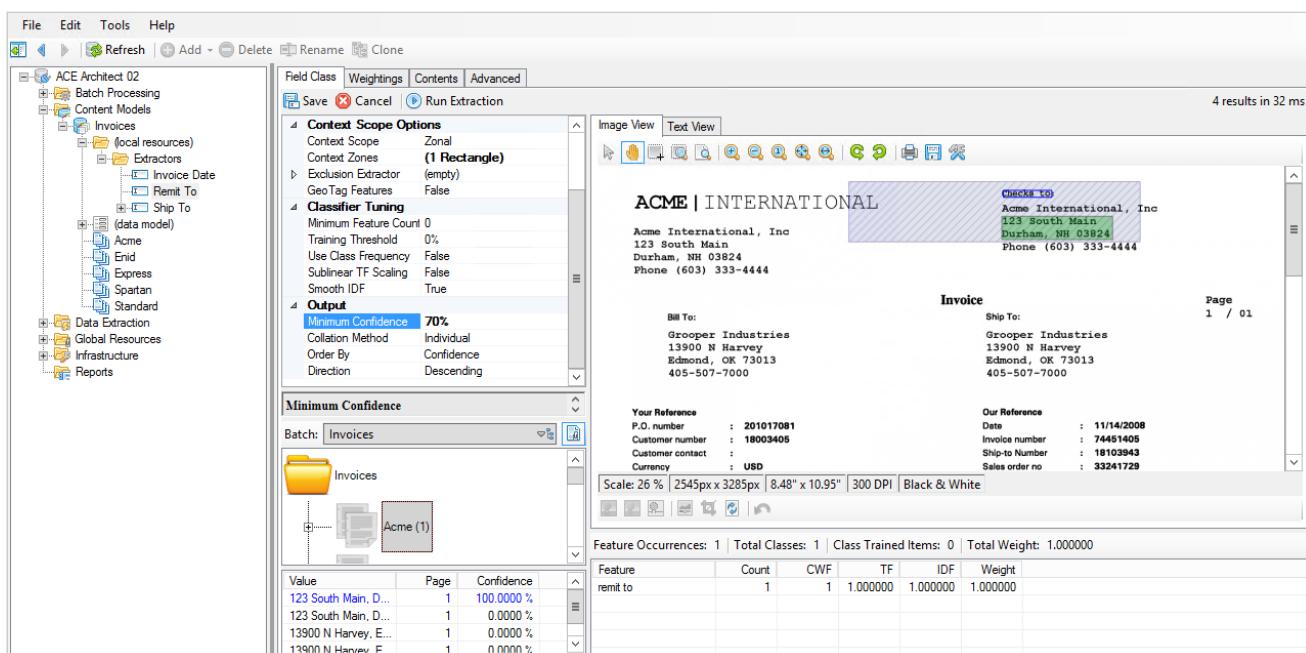
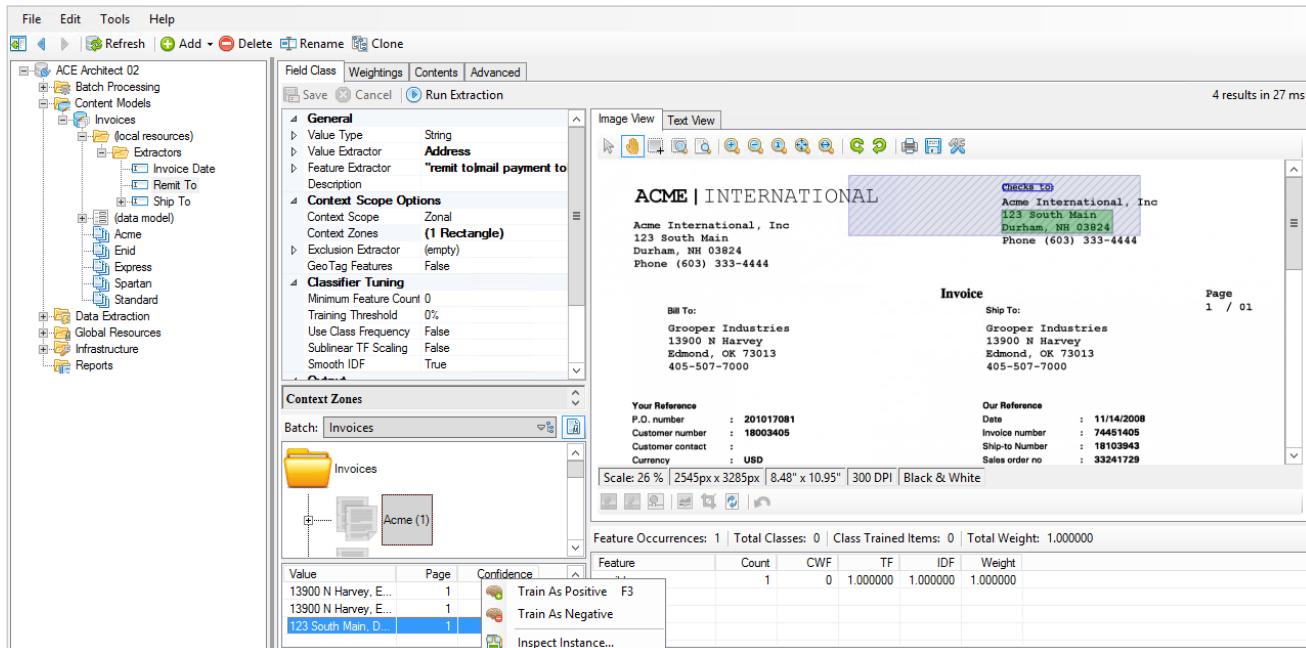
1. Click to edit the Context Zones property
2. In the Context Zones window, delete one of the zones
3. Edit the dimensions of the remaining zone:
 - a. Left: -2.55
 - b. Top: -0.55
 - c. Right: 1
 - d. Bottom: 0.15

4. Click OK



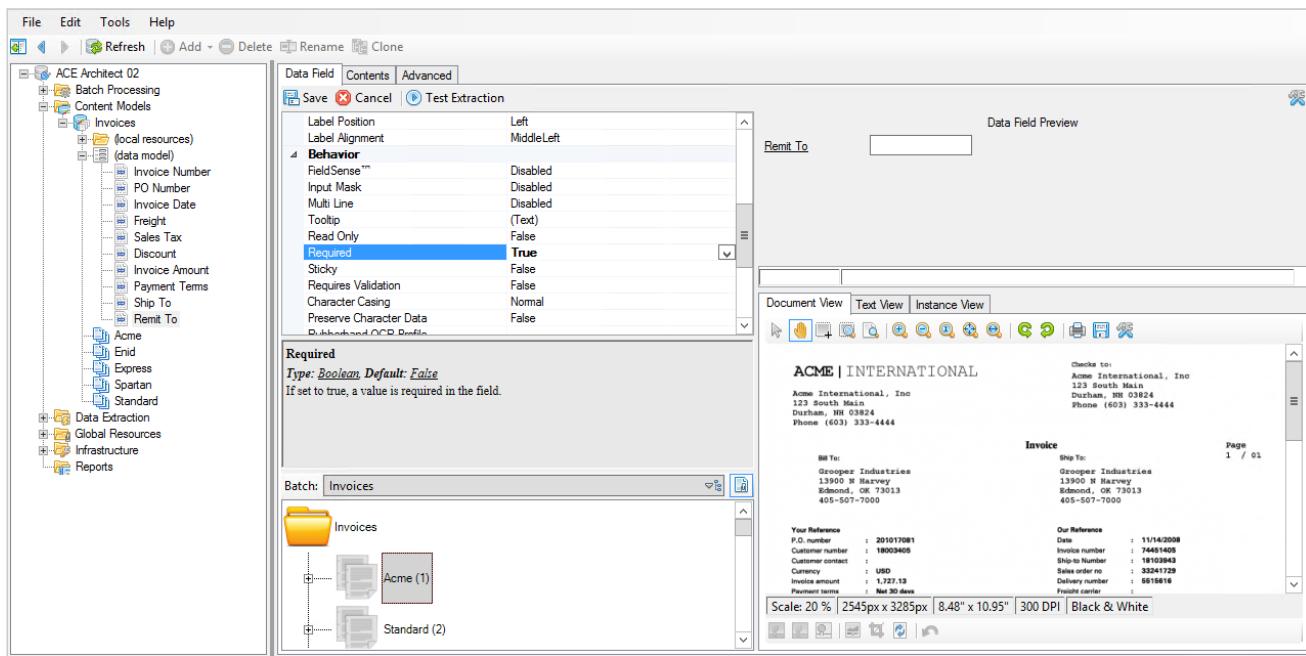
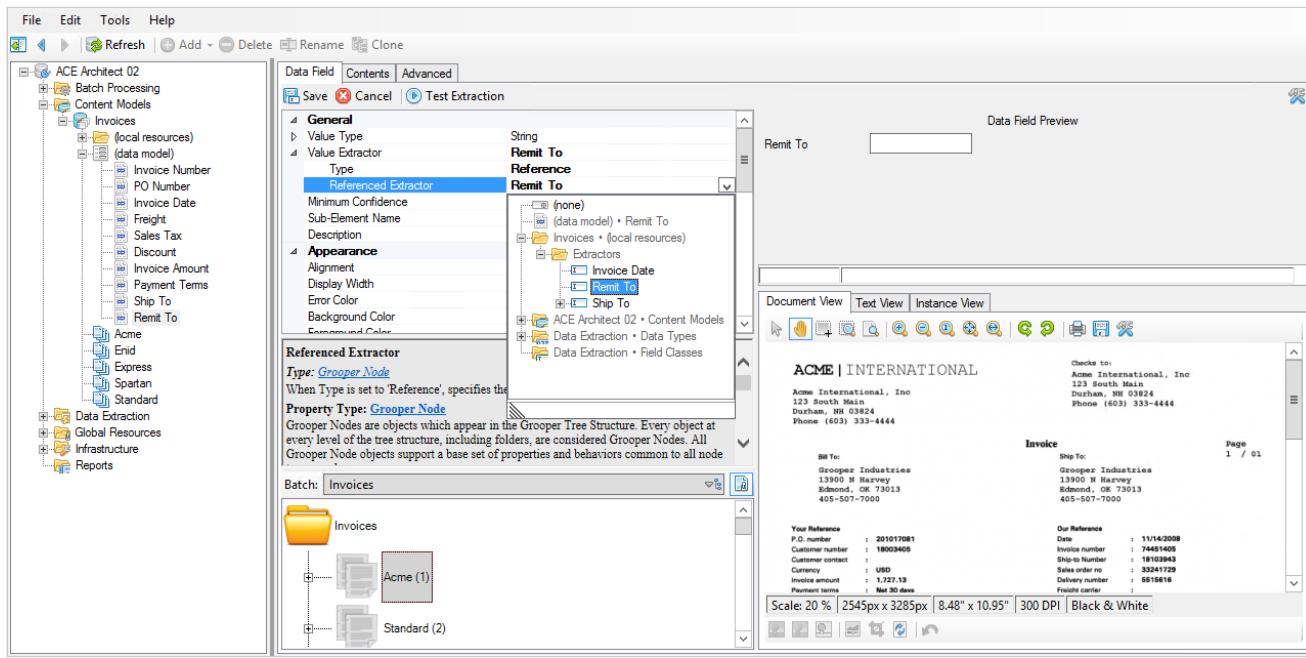
Step

1. Select Acme (1)
2. In the candidate list, find the value that has features that can be trained
3. Right click on that value and select Train As Positive
4. Set the Minimum Confidence property to **70%**



Step

1. In the Remit To Data Field, set the Value Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > Remit To Field Class.
3. In the Behavior section, set the Required property to **True**



Feature List

Step

1. Navigate to (root) > Batch Processing > Batches > Test > Invoices
2. Click on the Batch Viewer tab
3. Collapse the Node Tree using the button in the toolbar
4. Open Notepad.exe alongside Grooper Design Studio

File Edit Tools Help

Batch Batch Viewer Contents Advanced

Save Cancel Clone To Production... Pause

General

- Priority: 3
- Current Step: -1
- Description: (none)

Batch Information

- Created: 10/17/2018 4:42 PM
- Created By: BISvdekk
- Current Step No: -1
- Current Step Name: (none)
- Current Step Activity Name: (none)

Batch

A Batch is a hierarchy of folders and pages which serves as the fundamental unit of capture in Grooper.

Remarks

Batches can exist in one of two environments: Test or Production. Test

Processing Status Statistics Tasks Batch Viewer Processing History Event Log Details

Task Status

Pending In Progress Completed Errors

File Edit Tools Help

Batch Batch Viewer Contents Advanced

Batch

ACME | INTERNATIONAL

Checks to:

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	: 122565E90301034108

Invoice details

Scale: 28 % | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

File Edit Tools Help

Batch Batch Viewer Contents Advanced

Batch

ACME | INTERNATIONAL

Checks to:

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

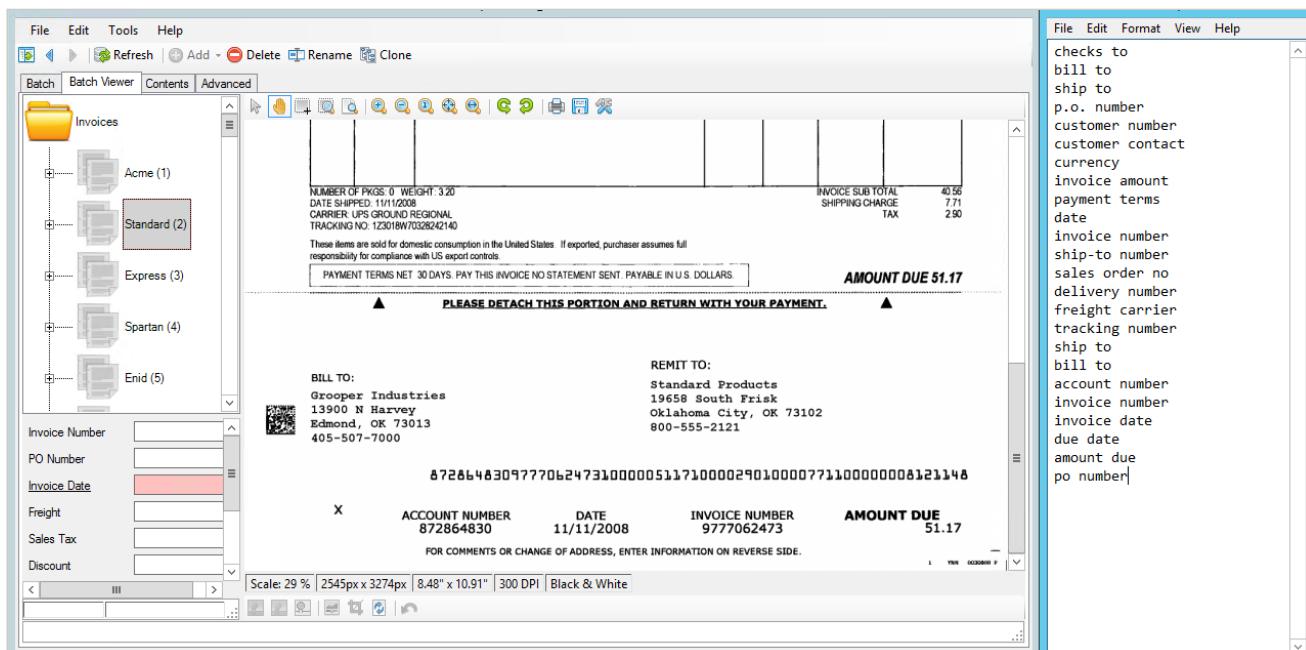
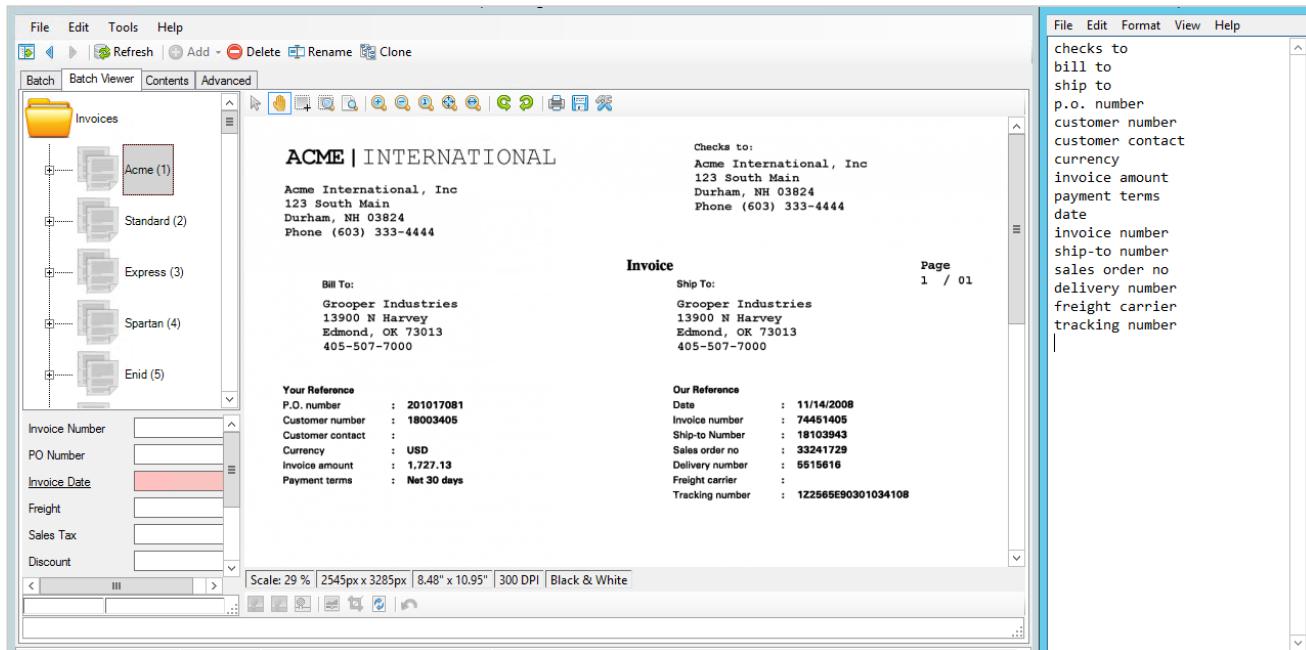
Date	: 11/14/2008
Invoice number	: 74451405
Ship-to Number	: 18103943
Sales order no	: 33241729
Delivery number	: 5515616
Freight carrier	:
Tracking number	: 122565E90301034108

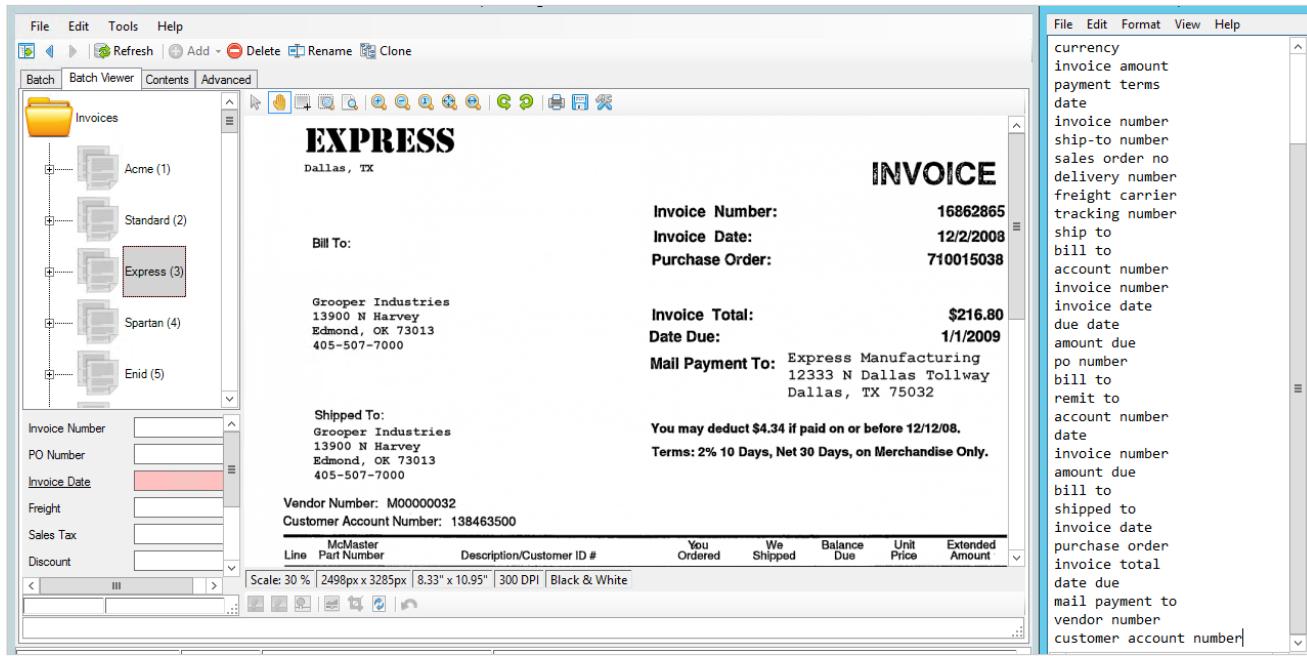
Invoice details

Scale: 29 % | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

Step

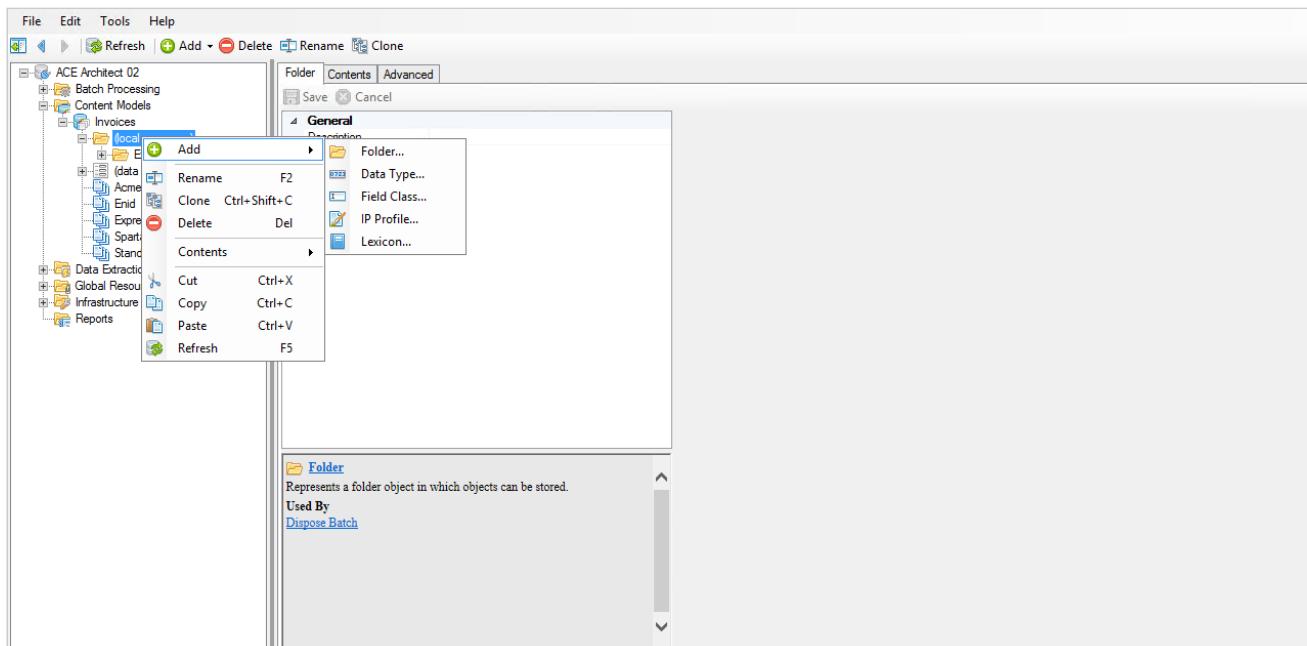
1. Select Acme (1)
2. In Notepad, begin typing a list of all words and phrases used to identify invoice items
3. Do this for the remaining four vendors
4. When complete, copy everything in Notepad to your clipboard

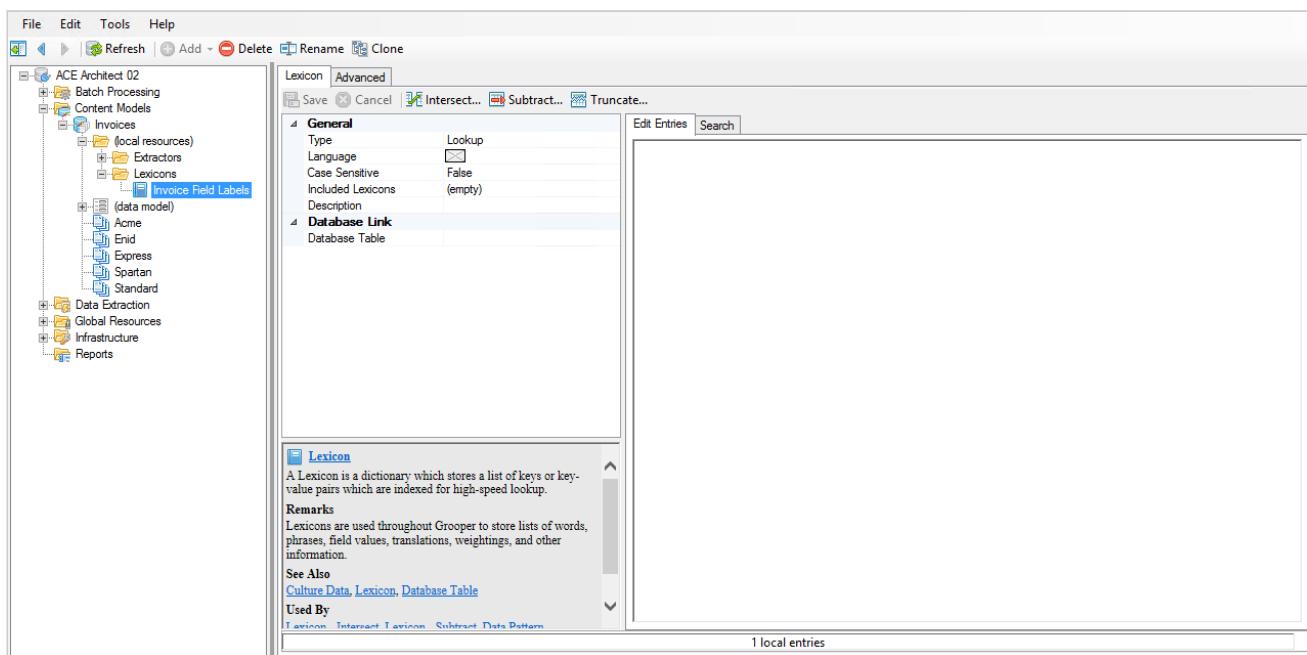
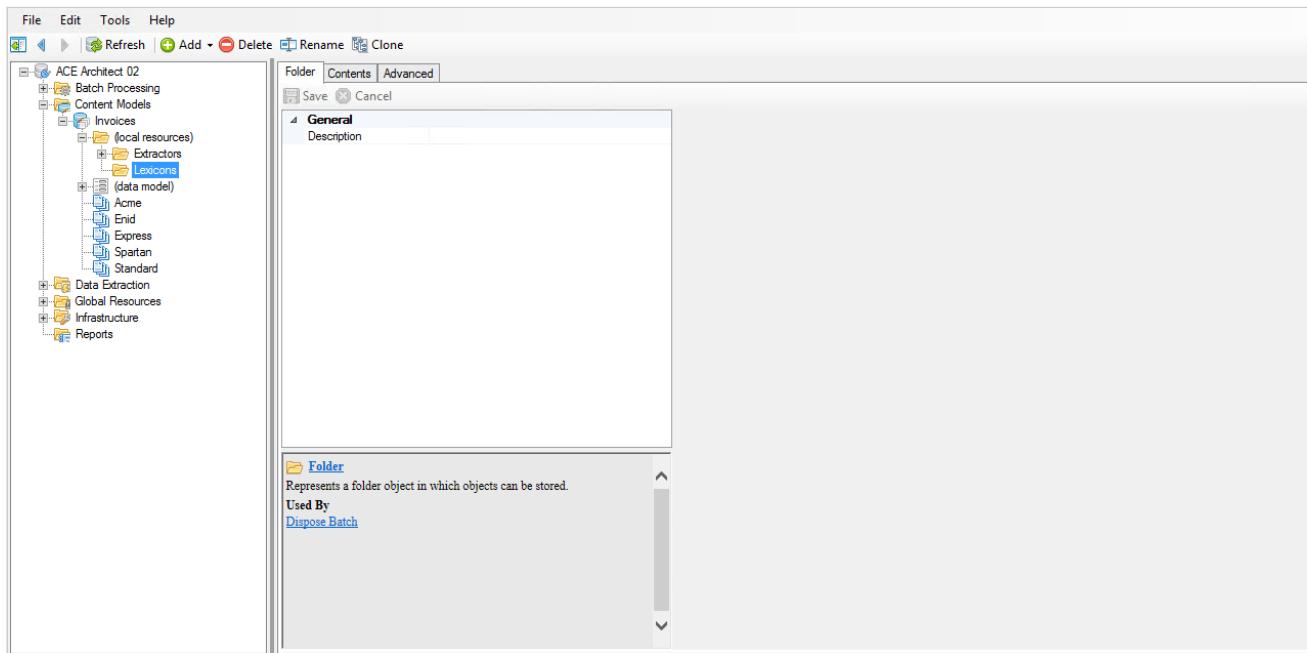




Step

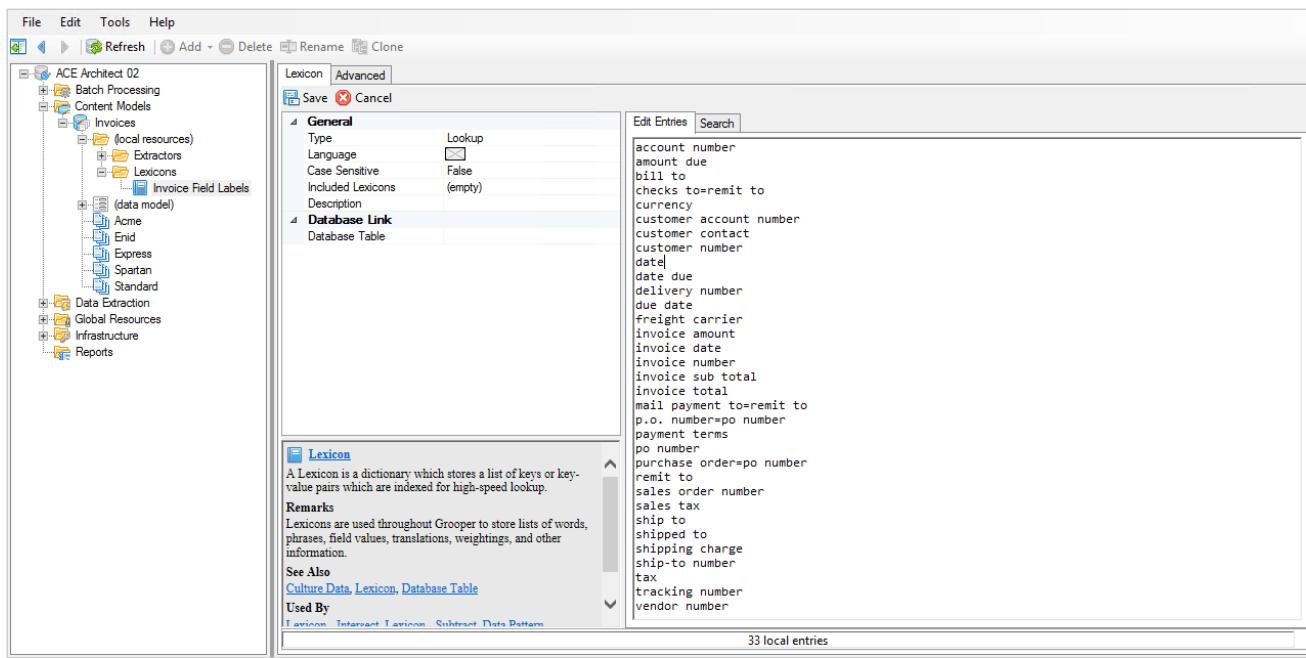
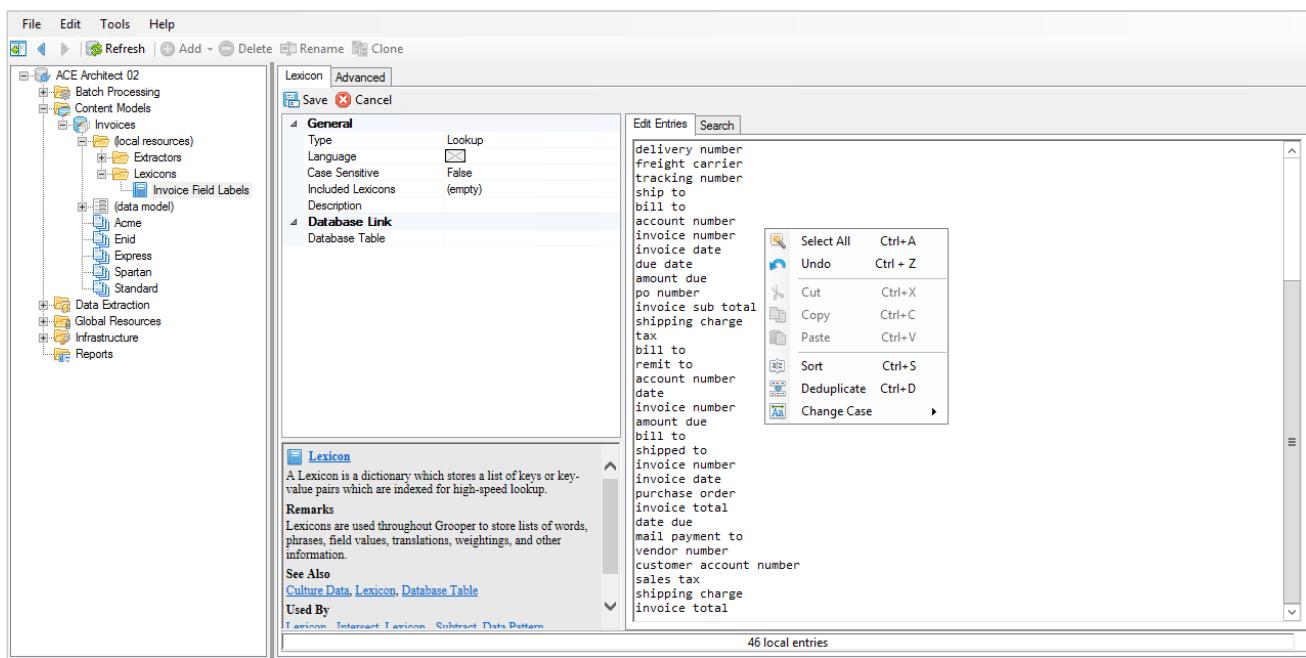
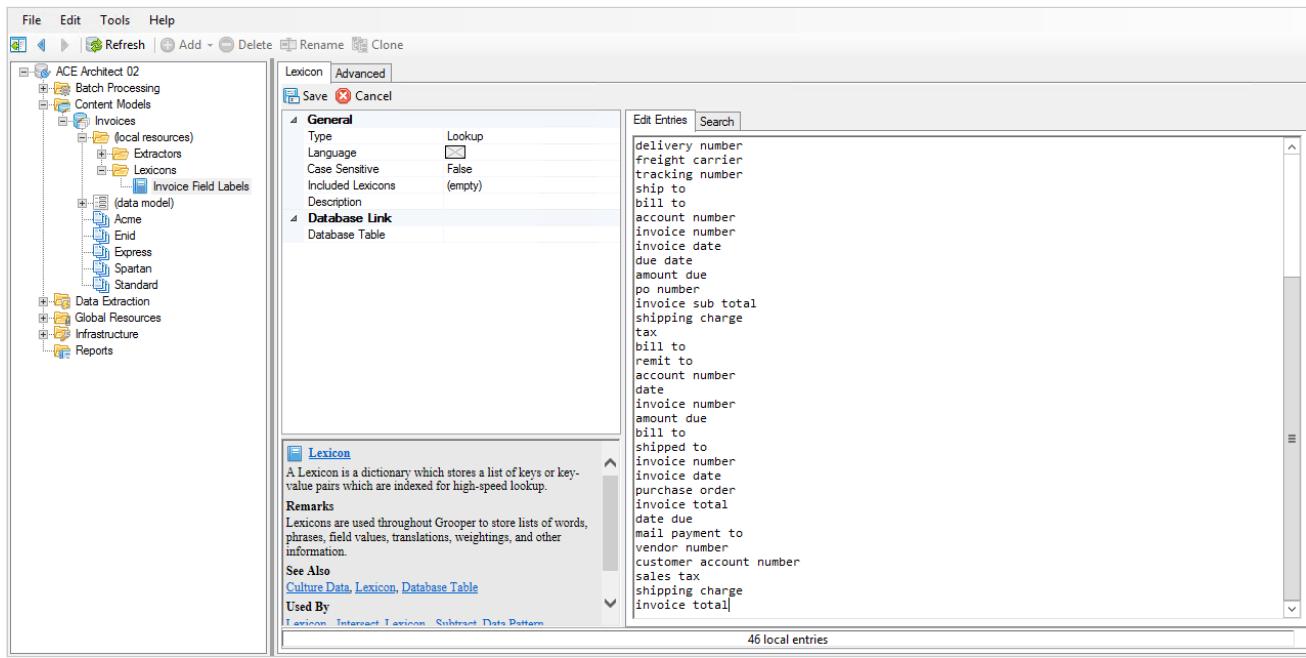
1. Restore the default Grooper Design Studio view
2. Navigate to Content Models > Invoices > (local resources)
3. Right click and select Add > Folder...
4. Name it **Lexicons**
5. Right click on the Lexicons folder and select Add > Lexicon...
6. Name it **Invoice Field Labels**





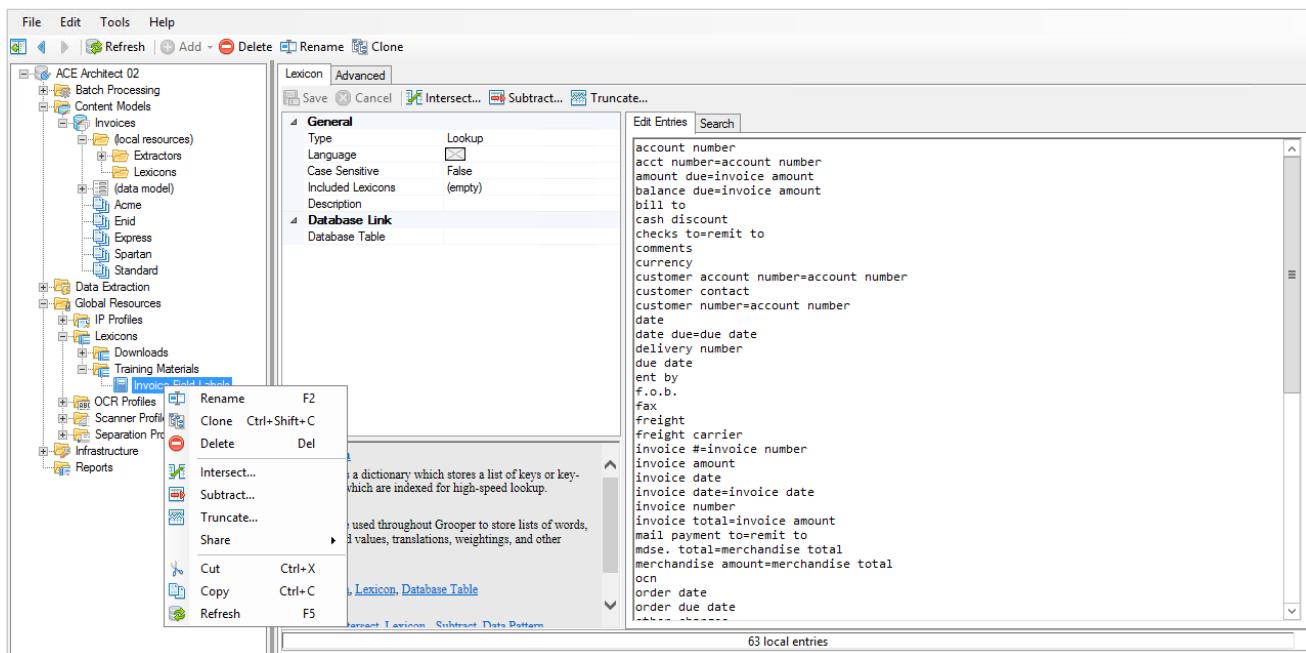
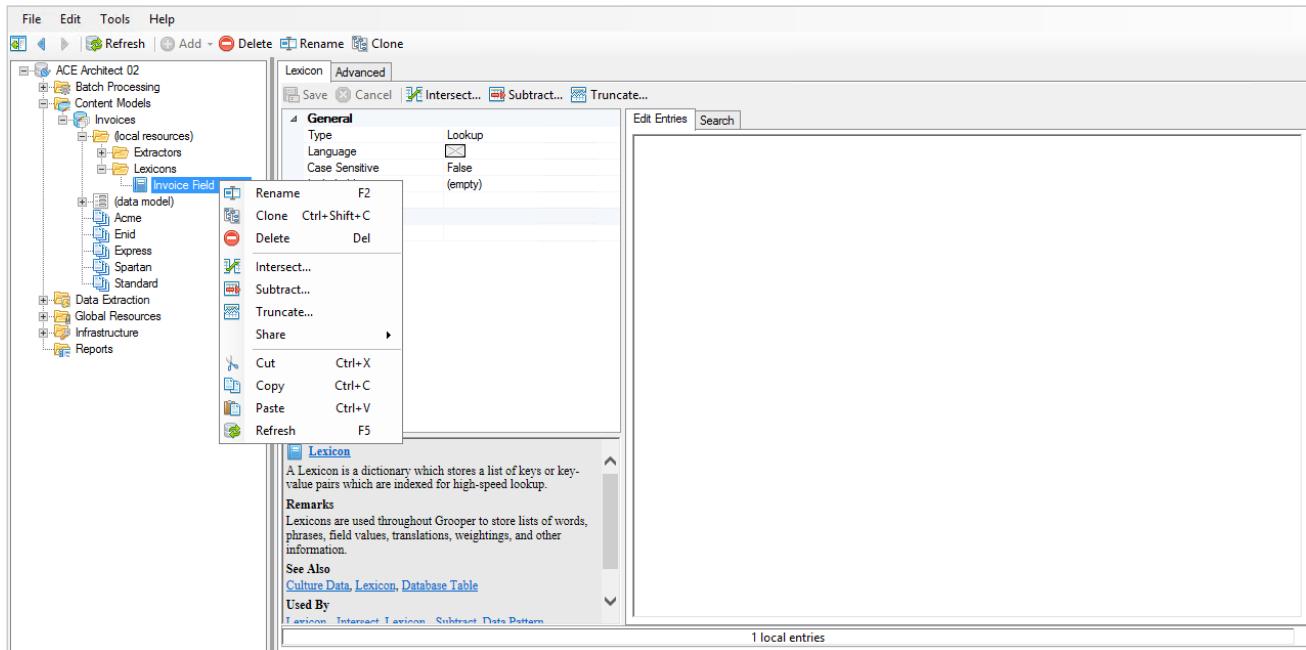
Step

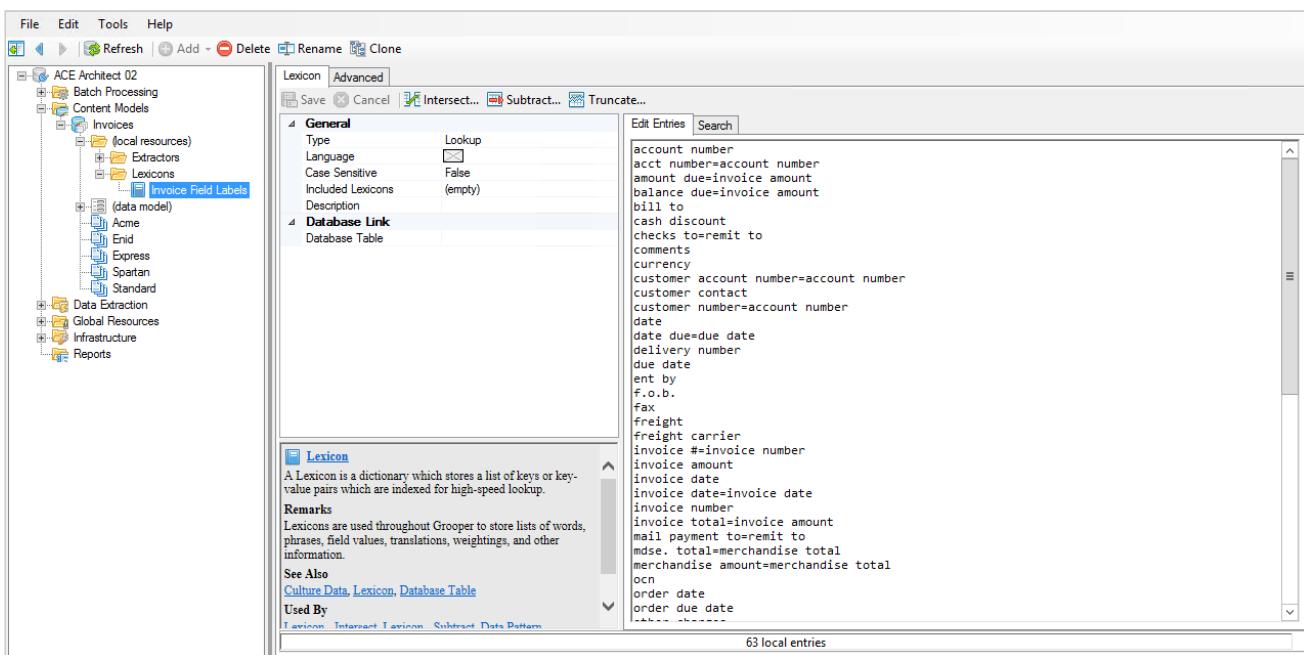
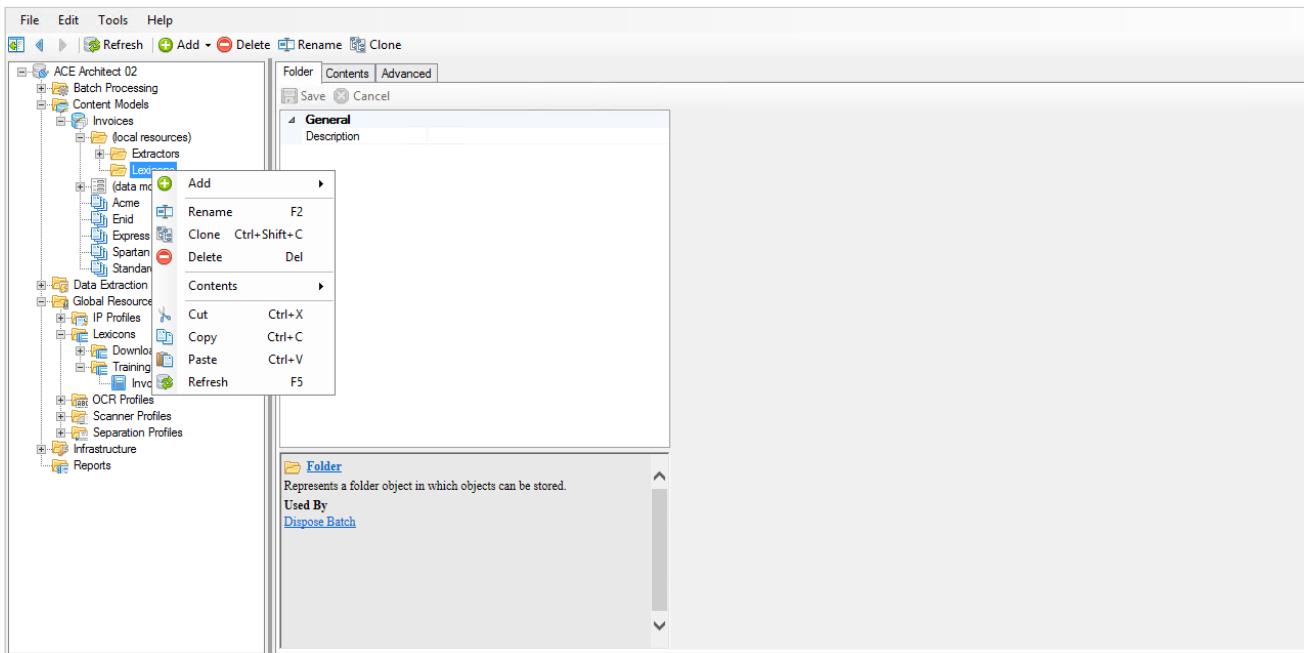
1. Paste your list of field labels in the Edit Entries editor
2. Right click in the Edit Entries editor and select Sort
3. Right click in the Edit Entries editor and select Deduplicate
4. Normalize synonymous phrases by setting an = sign between the original phrase and the desired transformed phrase



> Step

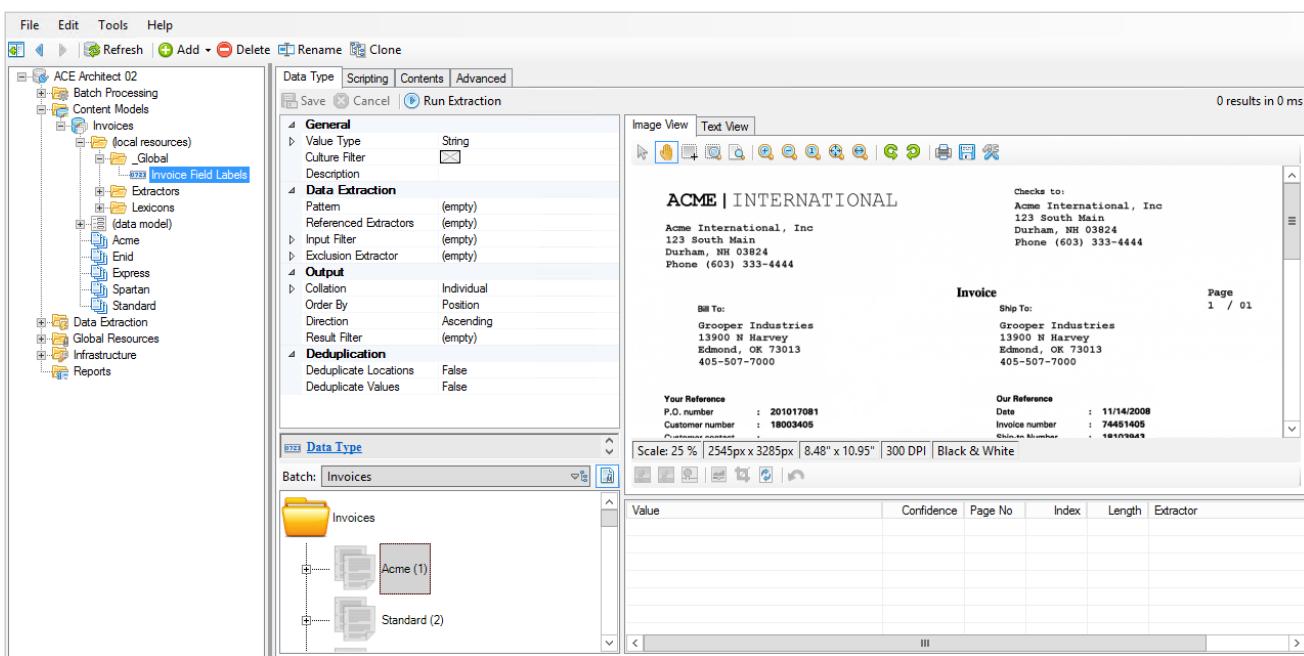
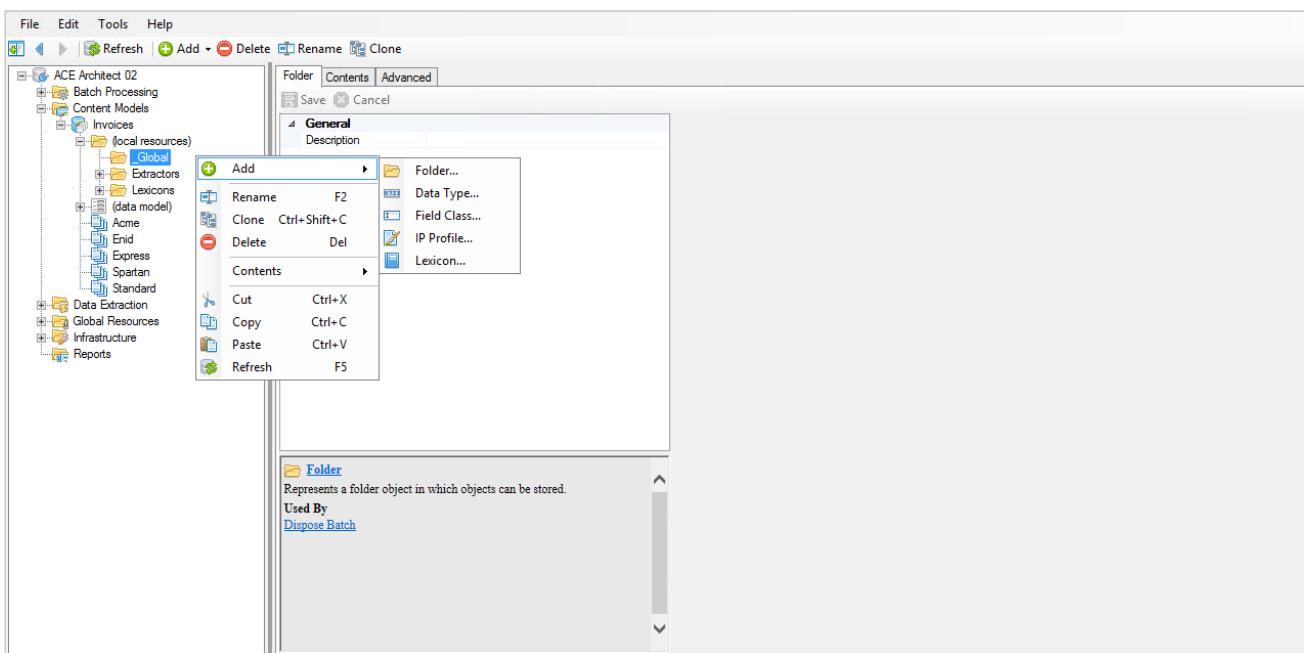
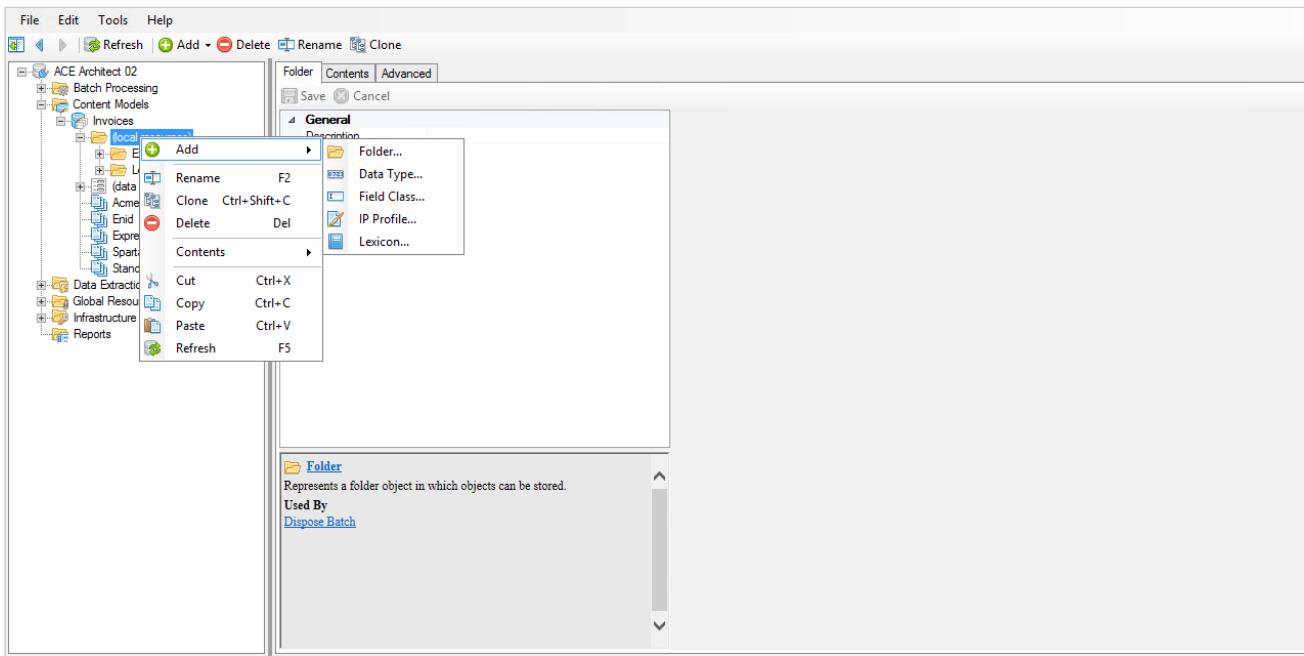
1. Delete this Lexicon
2. Navigate to (root) > Global Resources > Lexicons > Training Materials > Invoice Field Labels
3. Right click and select Copy
4. Right click on the Invoices > (local resources) > Lexicons folder and select Paste

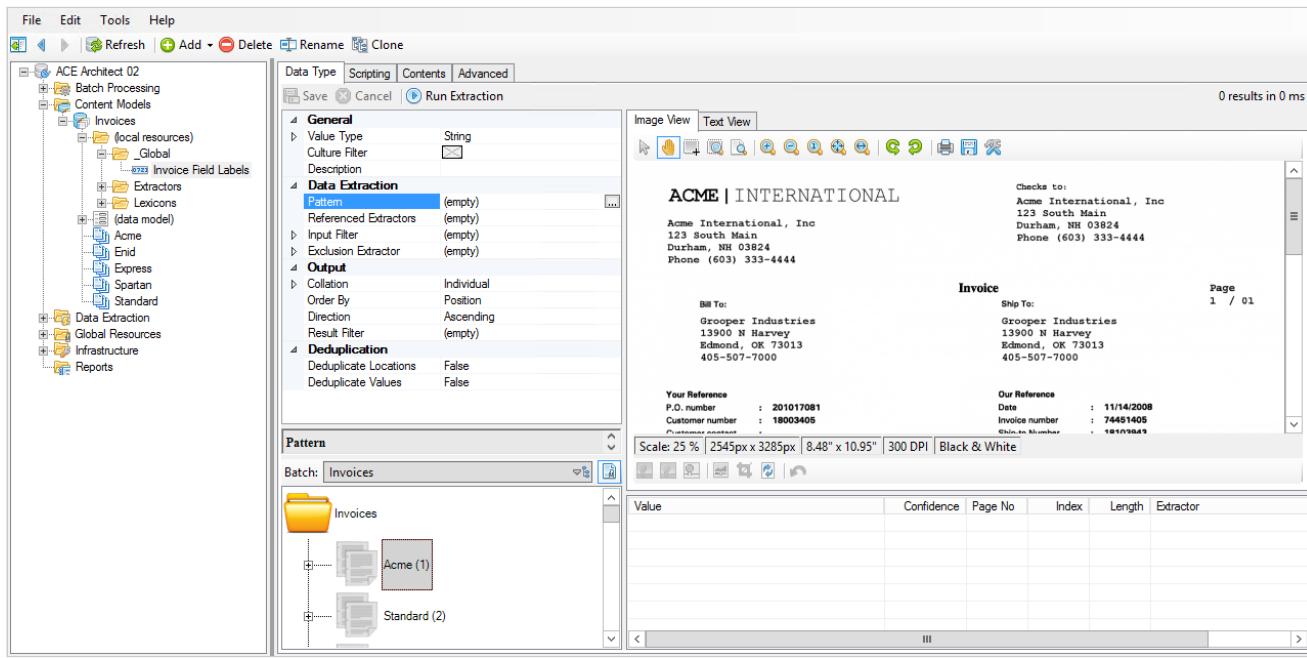




Step

1. Right click on the (local resources) folder and select Add > Folder...
2. Name it **_Global**
3. Right click on the _Global folder and select Add > Data Type...
4. Name it **Invoice Field Labels**
5. Click to edit the Pattern



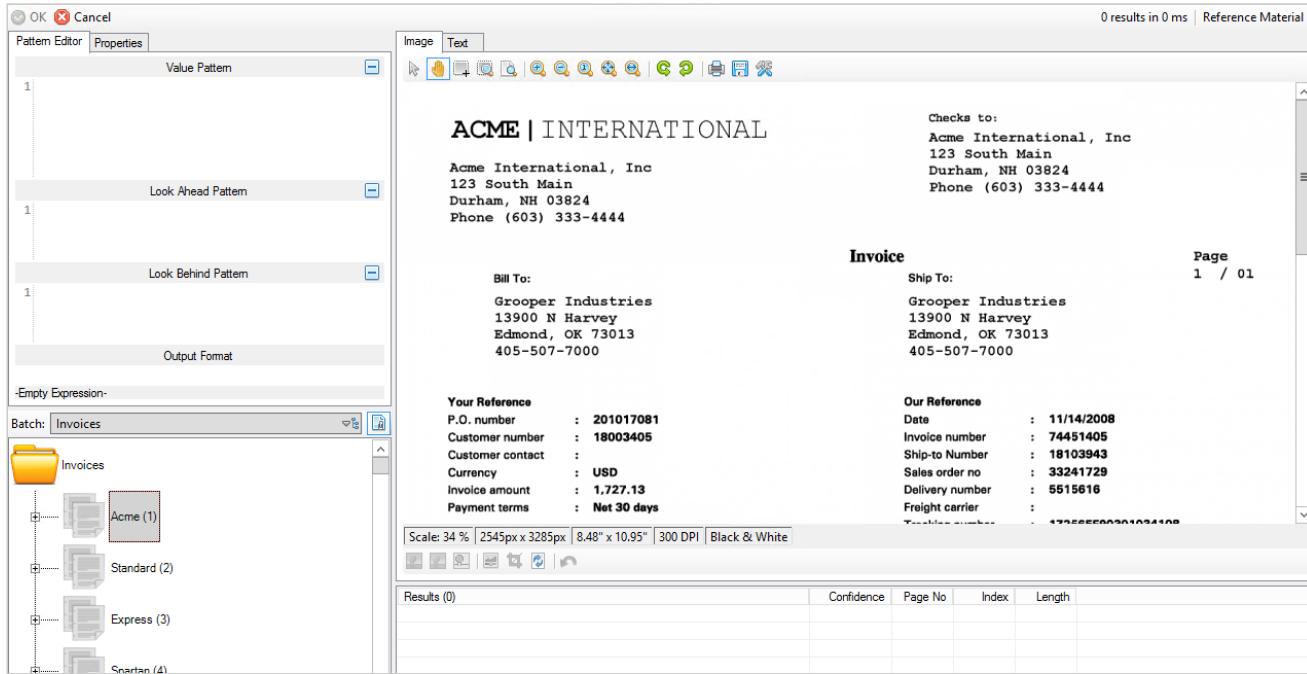


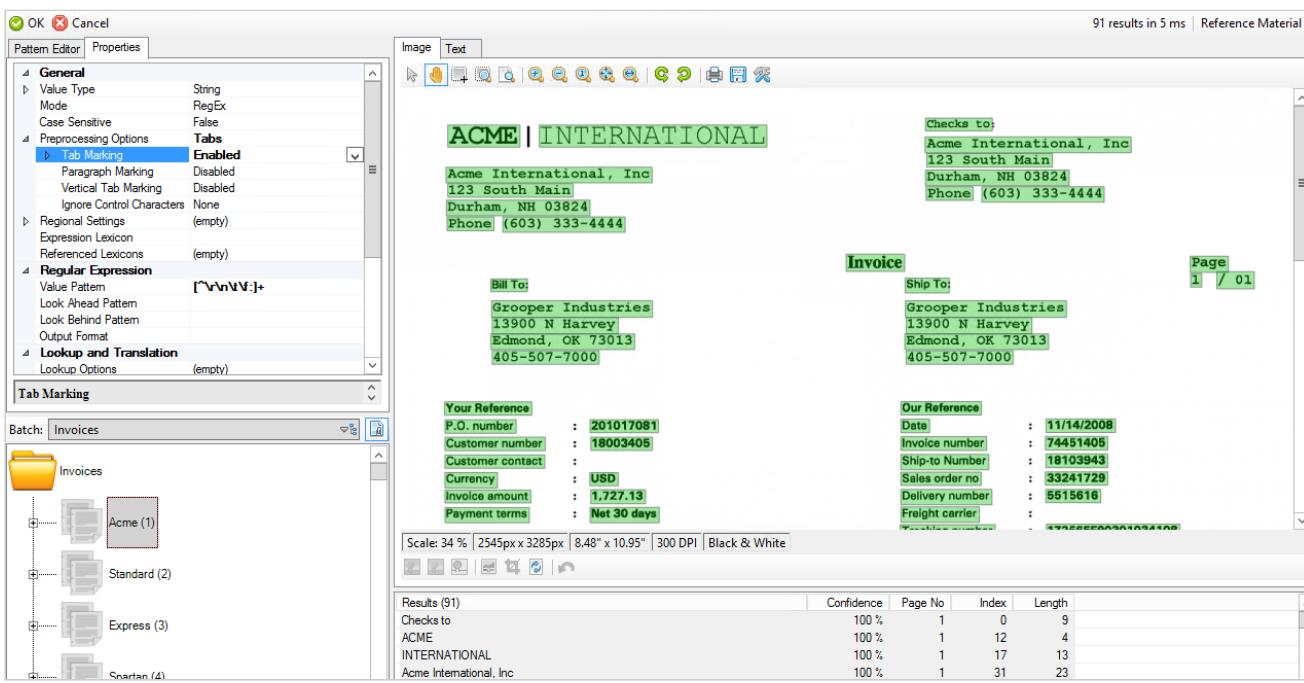
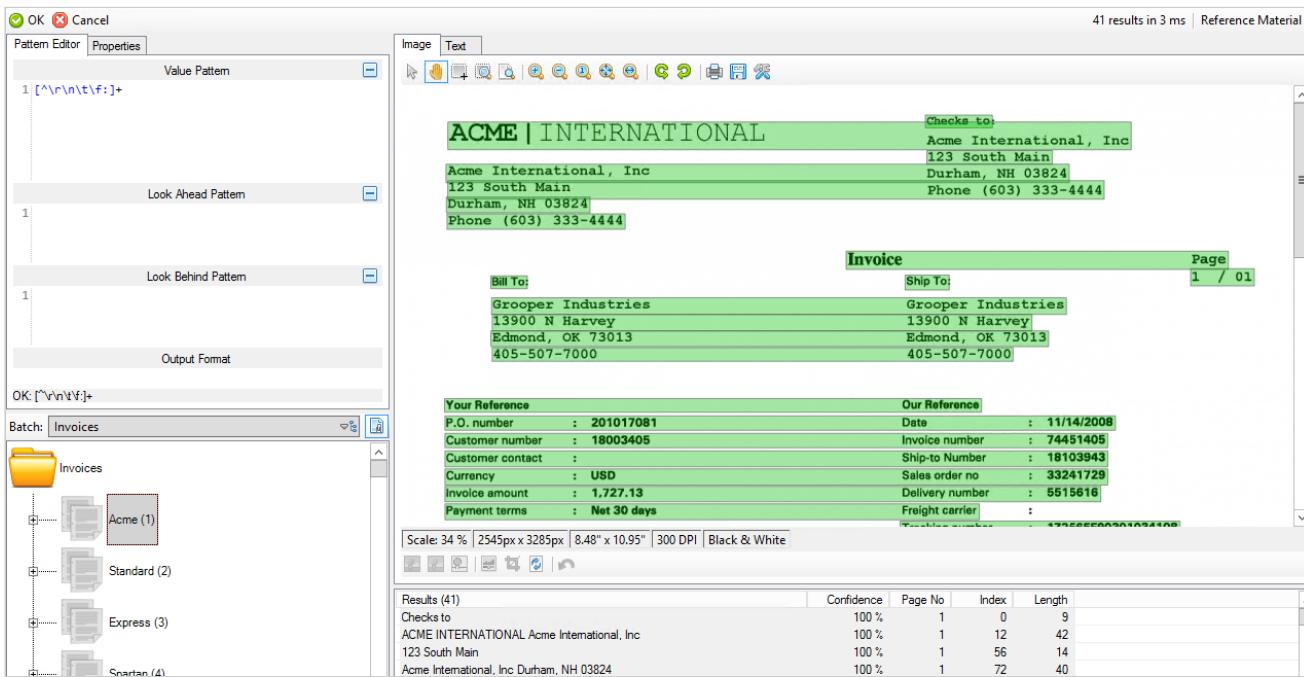
Step

1. In the Value Pattern editor, type:

```
[^\r\n\t\f:]+
```

2. Click on the Properties tab
3. Expand Preprocessing Options and set the Tab Marking property to Enabled





Step

- In the Lookup and Translation section, click to edit the Lookup Options
- In the Lookup Options window, expand Vocabulary
- Select the Vocabulary > Included Lexicons property dropdown and check the box next to the Invoices • (local resources) > Lexicons > Invoice Field Labels Lexicon
- Set the Enable Translation property to **True**
- Set the Fuzzy Match Similarity property to **90%**
- Click OK

Pattern Editor

General

- Value Type: String
- Mode: RegEx
- Case Sensitive: False
- Preprocessing Options: (empty)
- Regional Settings: Expression Lexicon
- Referenced Lexicons: (empty)

Regular Expression

- Value Pattern: `[^\v\n\v\v]+`
- Look Ahead Pattern:
- Look Behind Pattern:
- Output Format:

Lookup and Translation

- Lookup Options: (empty) (0 Group Lookup Options objects)
- Group Lookup Options: 0 entries
- Local Vocabulary Entries: 0 entries

Output Options

- Maximum Result Count: 0

Lookup Options

Batch: Invoices

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

ACME | INTERNATIONAL

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

Invoice

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

Page 1 / 01

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 :

Scale: 34 % | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

Results (91)

	Confidence	Page No	Index	Length
Checks to	100 %	1	0	9
ACME	100 %	1	12	4
INTERNATIONAL	100 %	1	17	13
Acme International, Inc	100 %	1	31	23

Pattern Editor

General

- Value Type: String
- Mode: RegEx
- Case Sensitive: False
- Preprocessing Options: Tabs (empty)
- Regional Settings: Expression Lexicon
- Referenced Lexicons: (empty)

Regular Expression

- Value Pattern: `[^\v\n\v\v]+`
- Look Ahead Pattern:
- Look Behind Pattern:
- Output Format:

Lookup and Translation

- Lookup Options: (empty) (0 Group Lookup Options objects)
- Group Lookup Options: 0 entries
- Local Vocabulary Entries: 0 entries

Output Options

- Maximum Result Count: 0

Lookup Options

Batch: Invoices

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

Lookup Options

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

Invoice

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

Page 1 / 01

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 :

Results (91)

	Confidence	Page No	Index	Length
Checks to	100 %	1	0	9
ACME	100 %	1	12	4
INTERNATIONAL	100 %	1	17	13
Acme International, Inc	100 %	1	31	23

Pattern Editor

General

- Value Type: String
- Mode: RegEx
- Case Sensitive: False
- Preprocessing Options: Tabs (empty)
- Regional Settings: Expression Lexicon
- Referenced Lexicons: (empty)

Regular Expression

- Value Pattern: `[^\v\n\v\v]+`
- Look Ahead Pattern:
- Look Behind Pattern:
- Output Format:

Lookup and Translation

- Lookup Options: (empty) (0 Group Lookup Options objects)
- Group Lookup Options: 0 entries
- Local Vocabulary Entries: 0 entries

Output Options

- Maximum Result Count: 0

Lookup Options

Batch: Invoices

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

Lookup Options

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

Invoice

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

Page 1 / 01

Your Reference

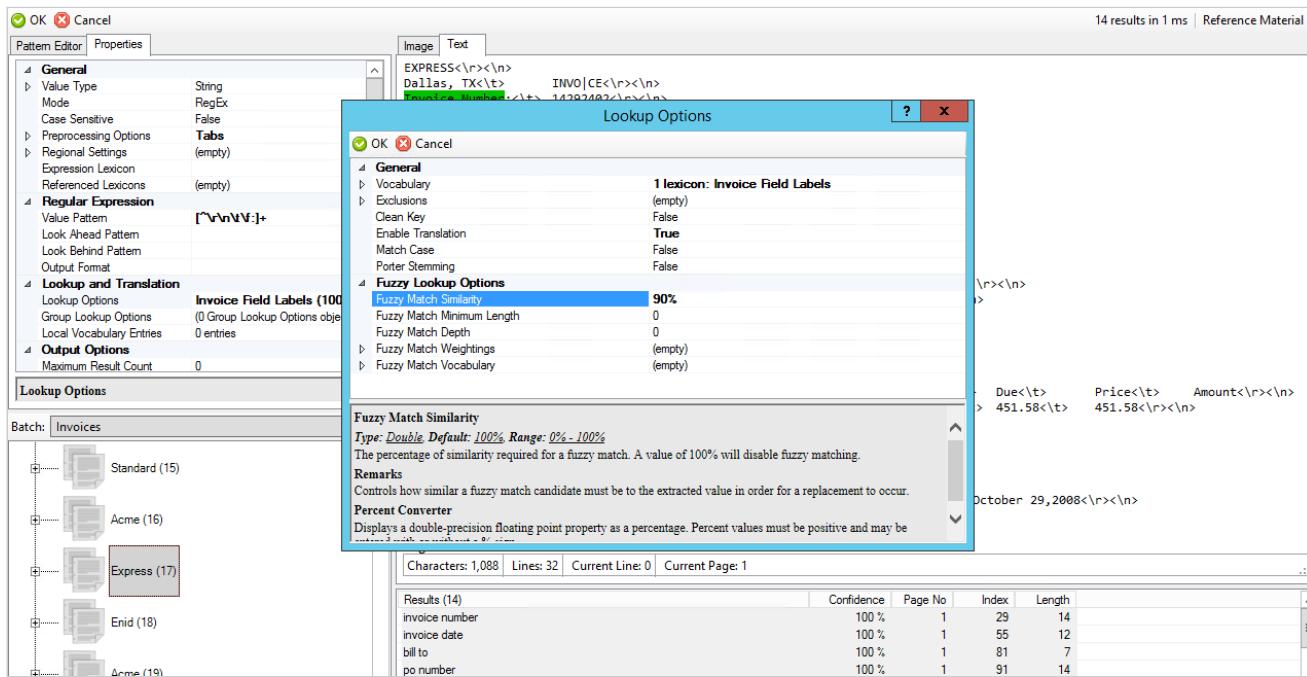
- P.O. number : 201017081
- Customer number : 18003405
- Customer contact :
- Currency : USD
- Invoice amount : 1,727.13
- Payment terms : Net 30 days

Our Reference

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

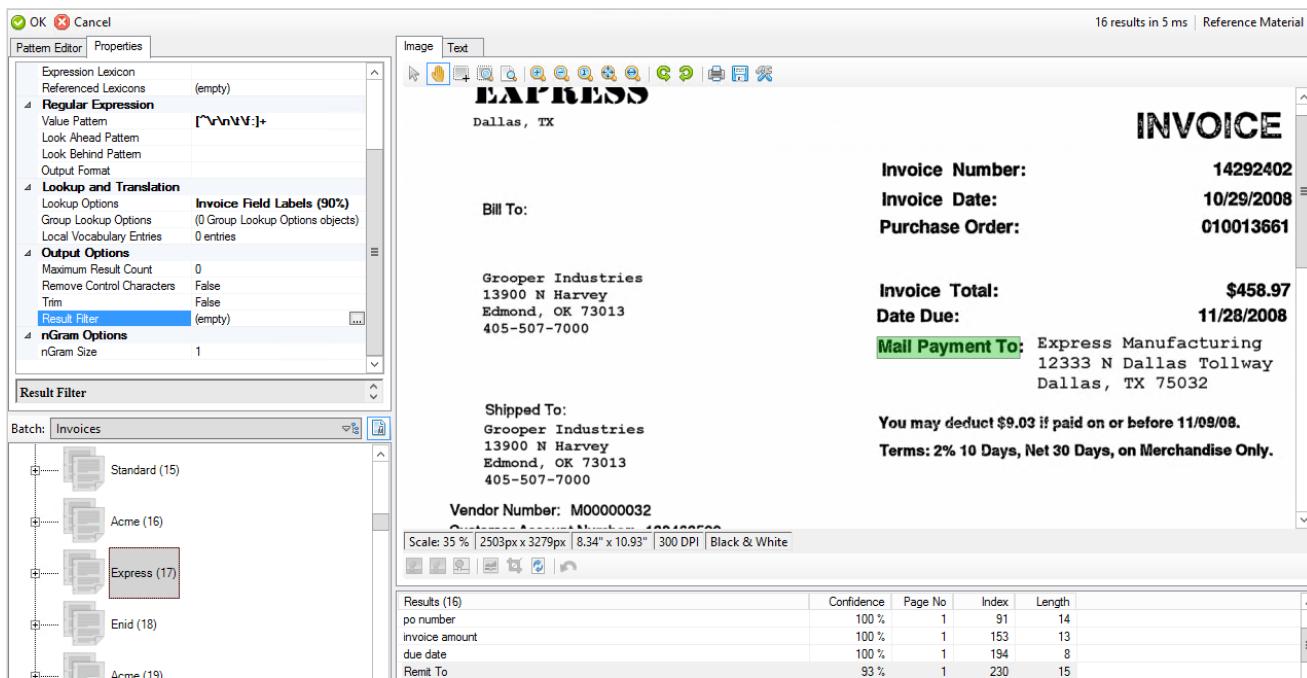
Results (91)

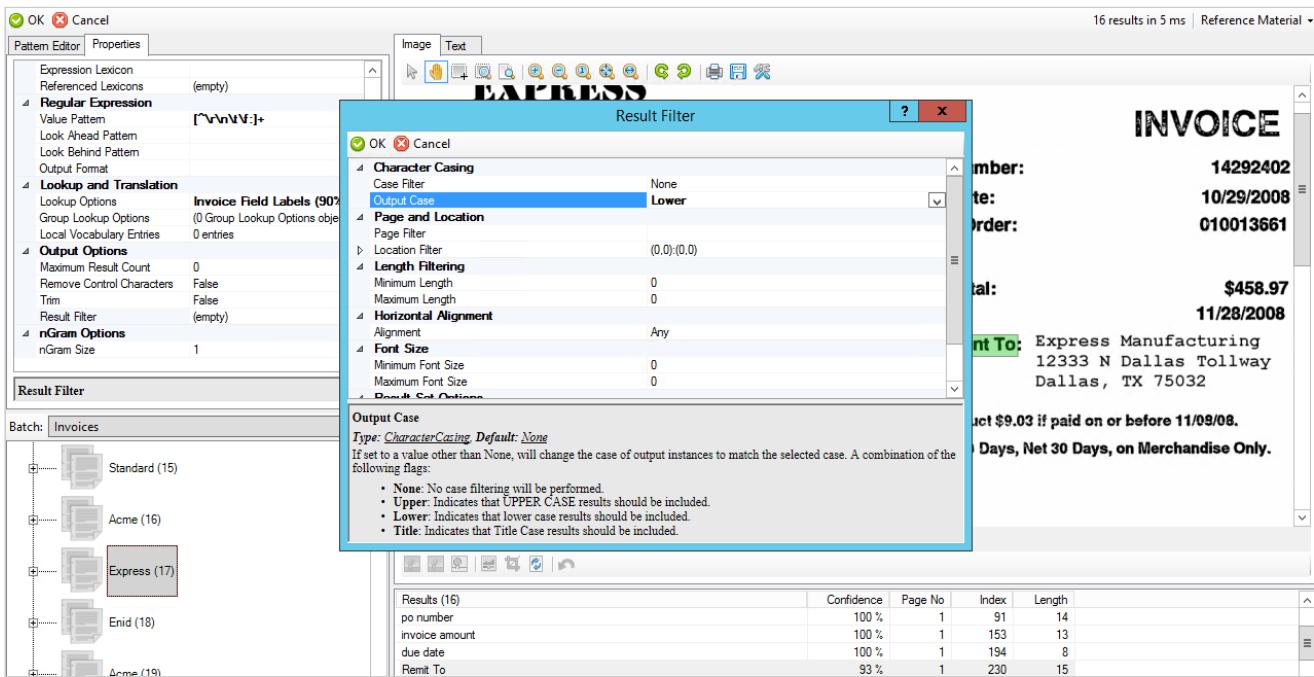
	Confidence	Page No	Index	Length
Checks to	100 %	1	0	9
ACME	100 %	1	12	4
INTERNATIONAL	100 %	1	17	13
Acme International, Inc	100 %	1	31	23



Step

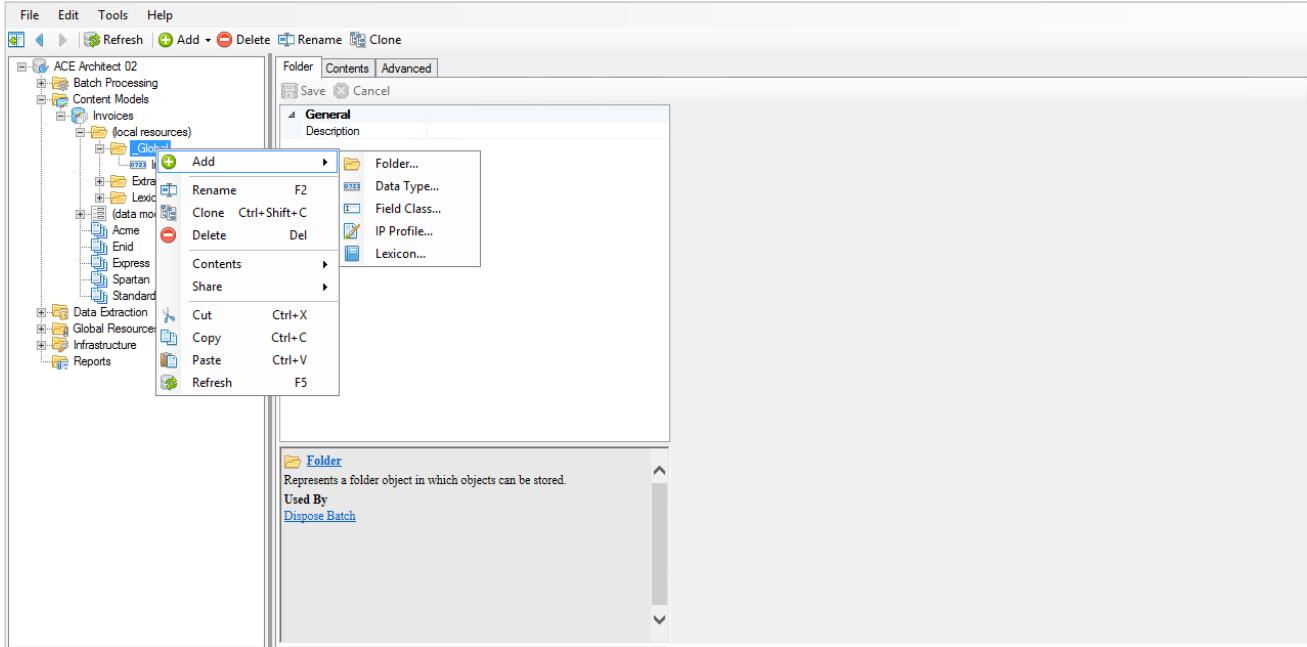
- In the Output Options section, click to edit Result Filter
- In the Result Filter window, set the Output Case property to **Lower**
- Click OK to close the Result Filter window
- Click OK to close the pattern editor window





Step

1. Right click on the (local resources) > _Global folder and select Add > Data Type...
2. Name it [Array] PO Number
3. Add two child Data Formats to the [Array] PO Number Data Type:
 - a. Purchase
 - b. Order Number



Step

1. Select the [Array] PO Number > Purchase Data Format

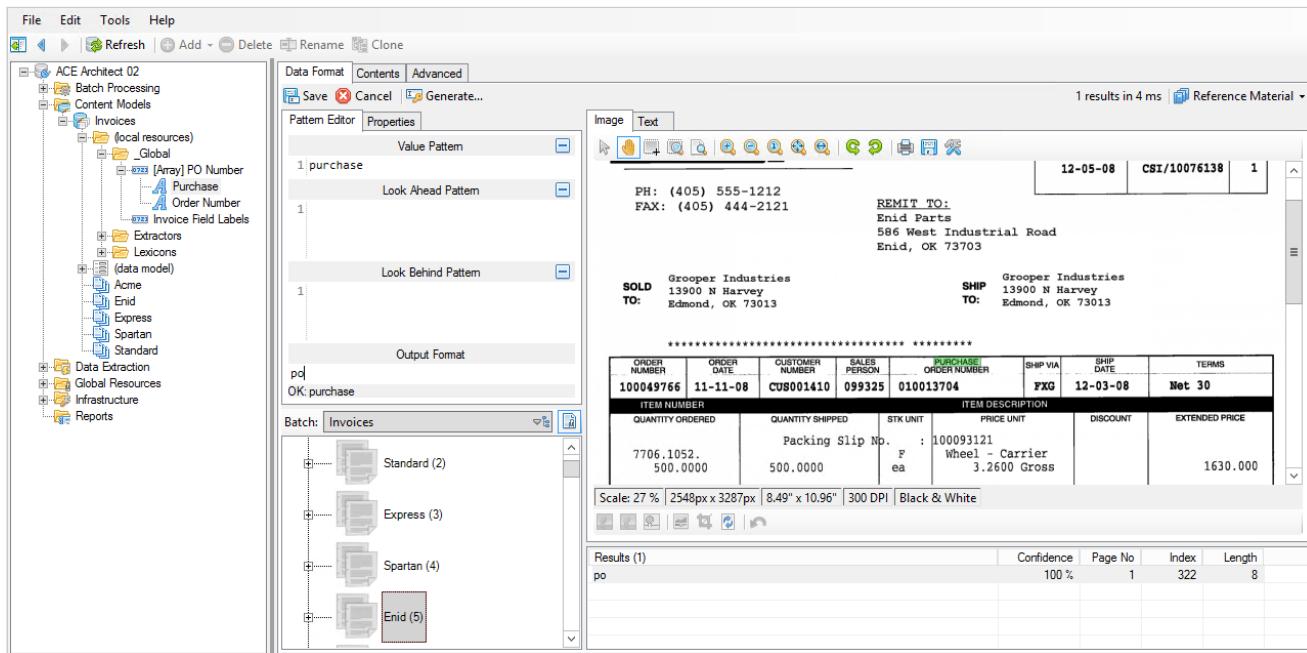
2. Select Enid (5)

3. In the Value Pattern editor, type:

purchase

4. In the Output Format editor, type:

po



Step

1. Select the [Array] PO Number > Order Number Data Format

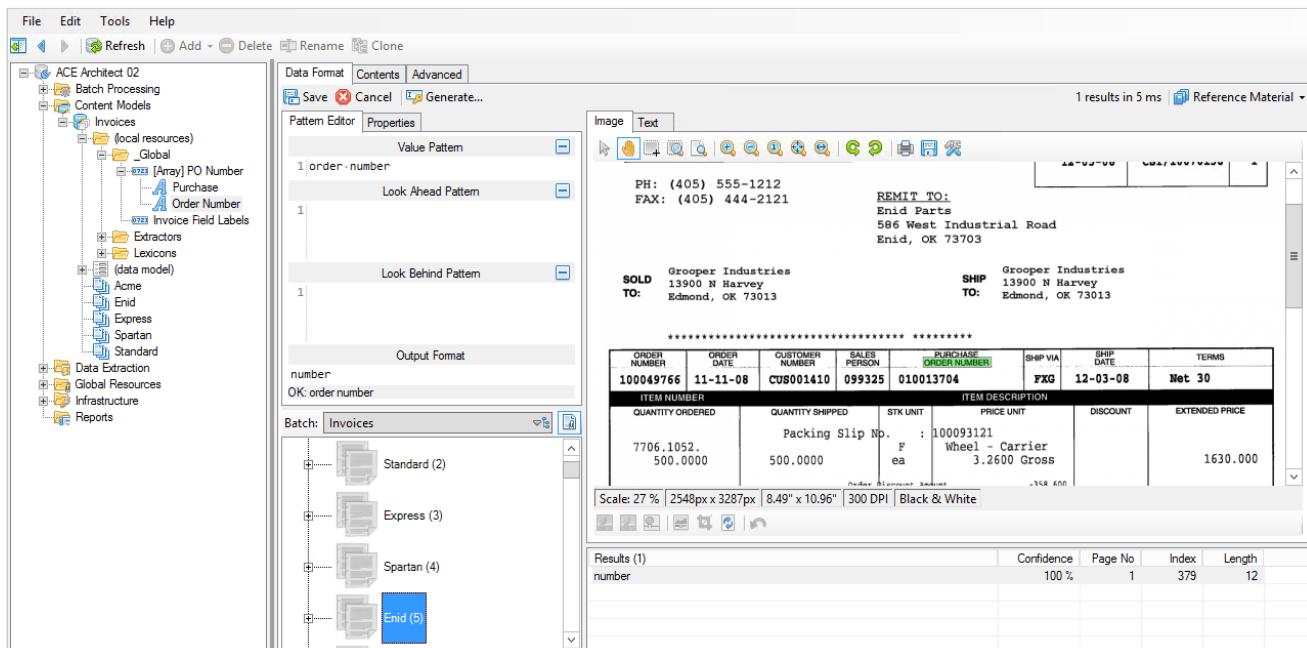
2. Select Enid (5)

3. In the Value Pattern editor, type:

```
order number
```

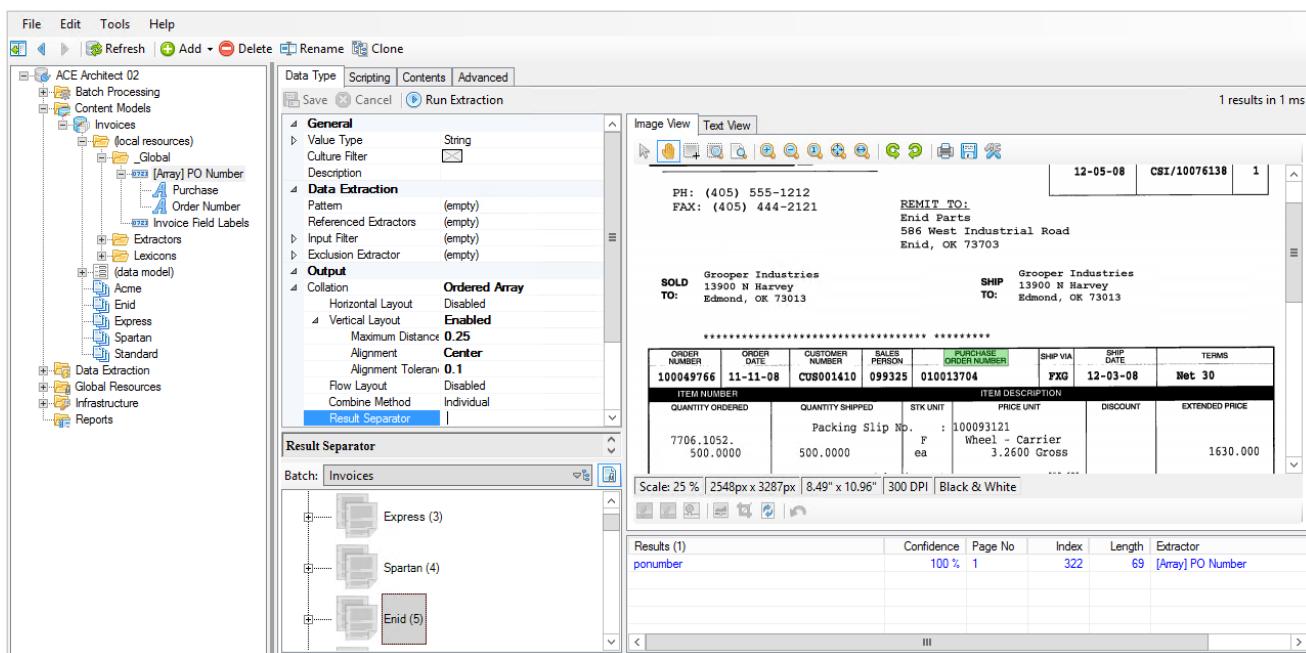
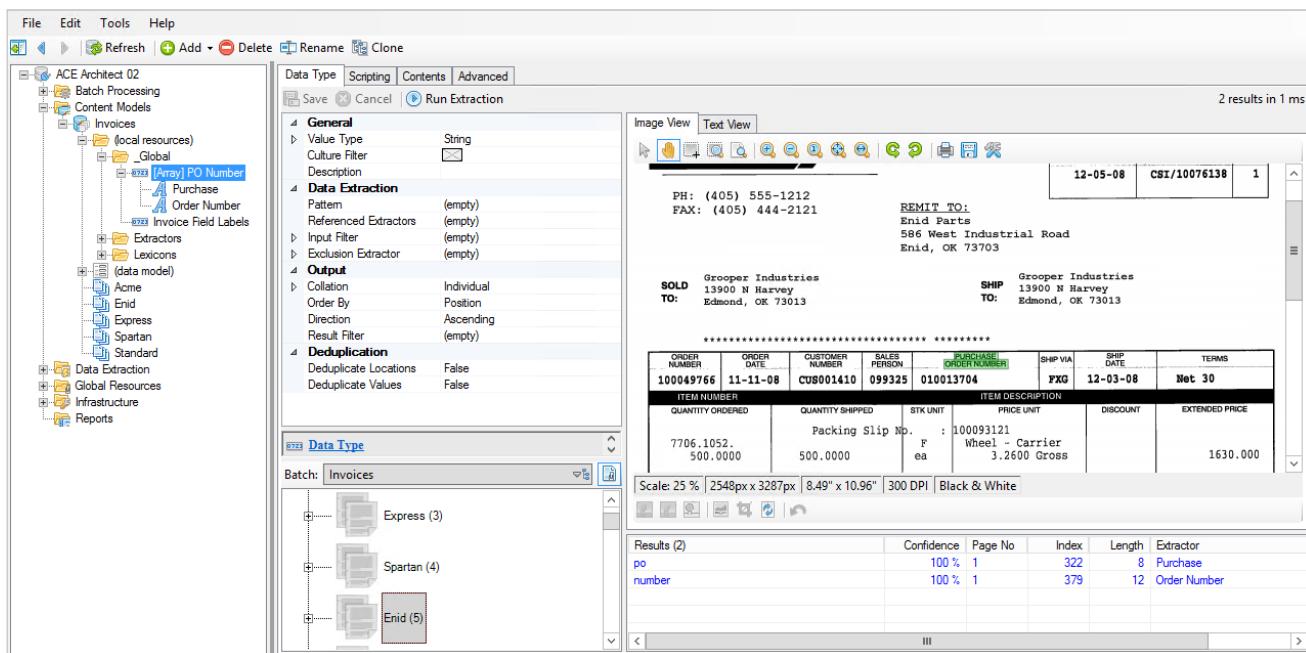
4. In the Output Format editor, type:

```
number
```



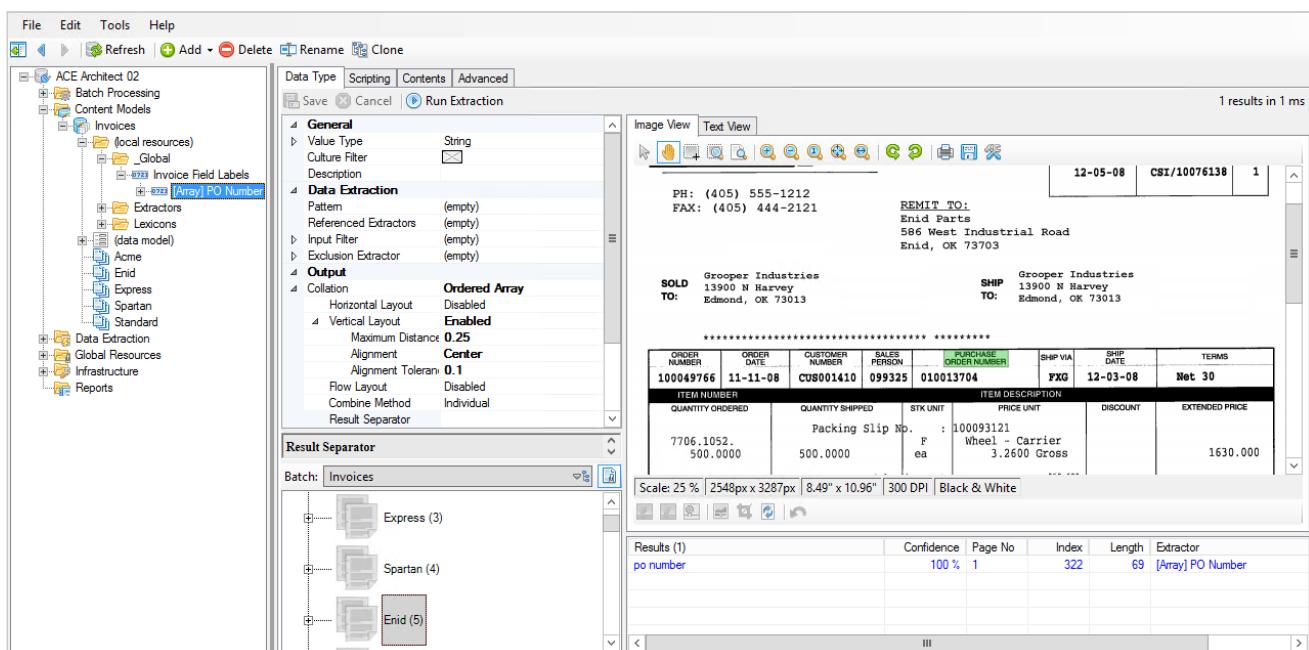
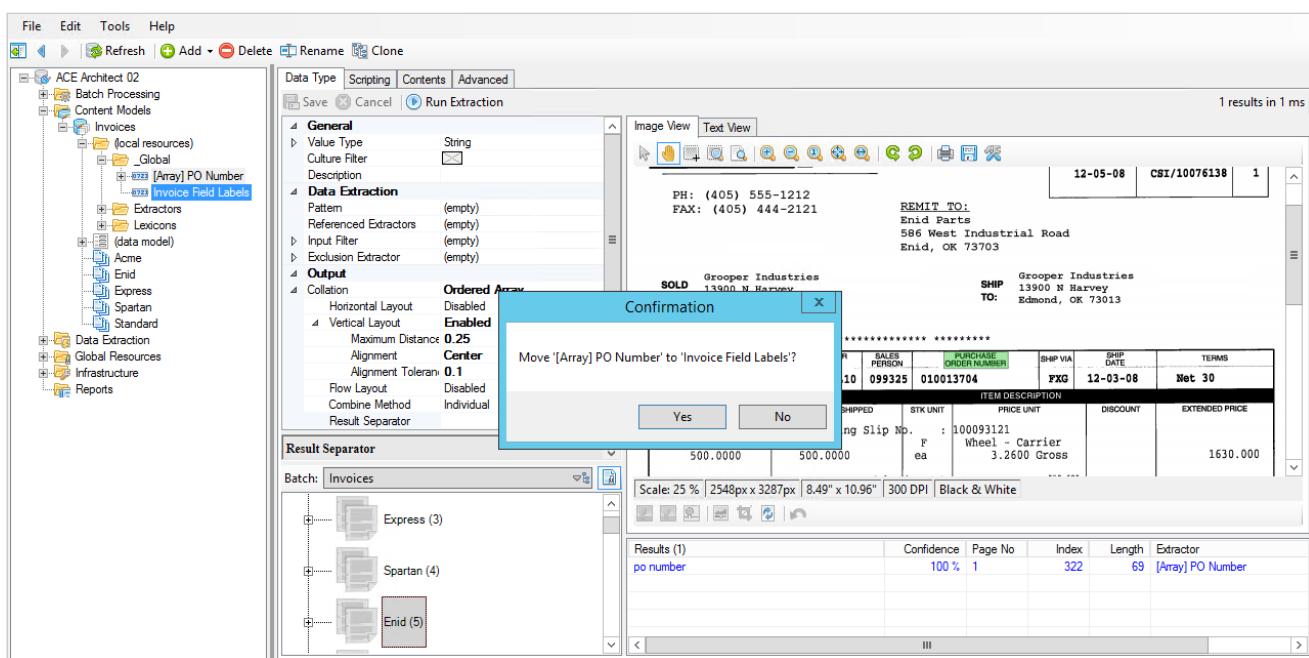
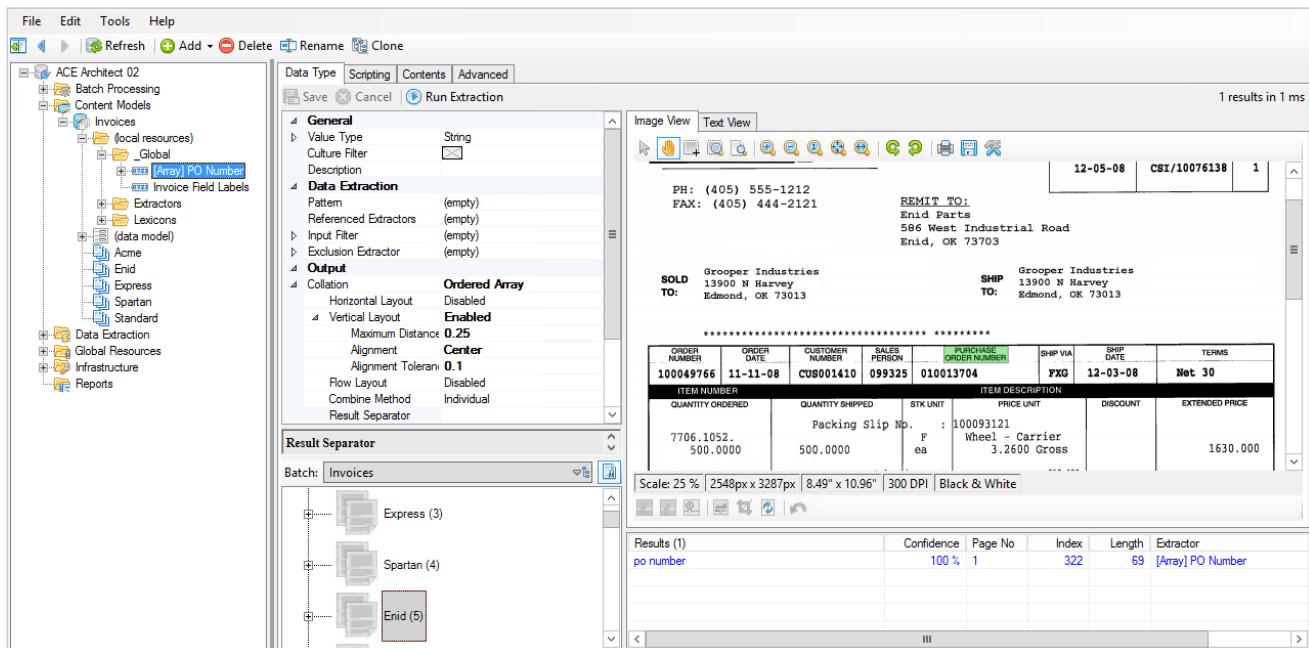
Step

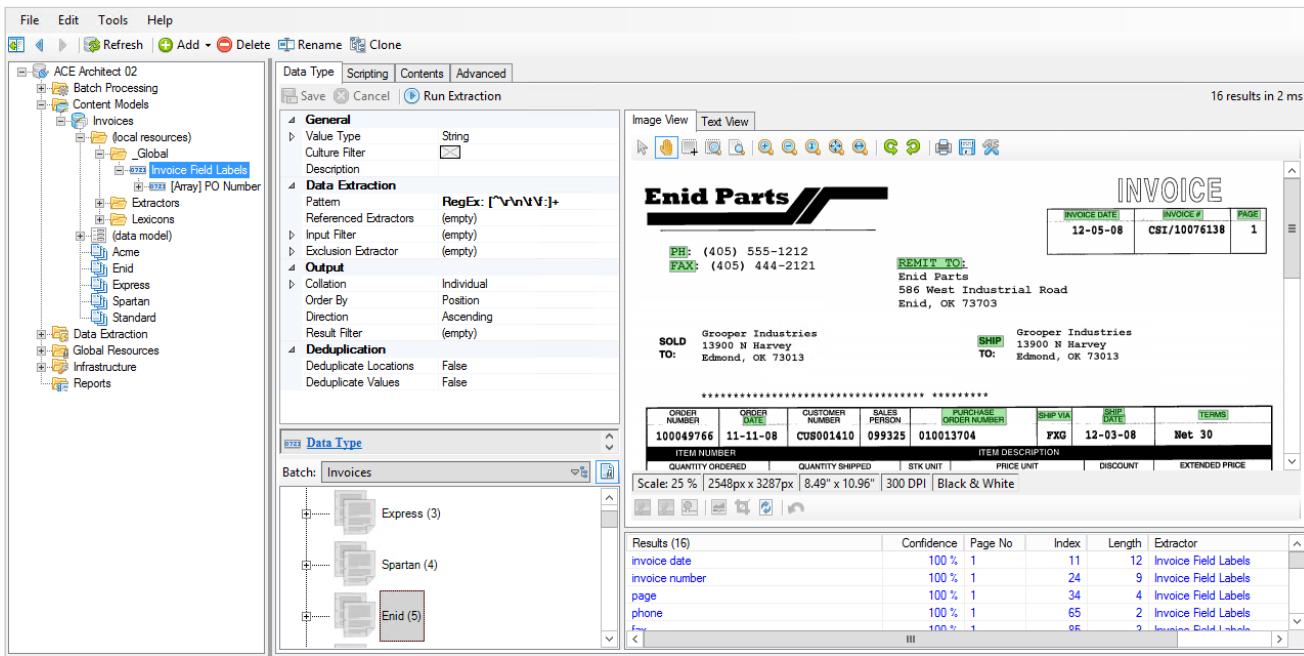
1. Select the [Array] PO Number Data Type
2. Set the Collation property to **Ordered Array**
3. Expand Collation and set the Vertical Layout property to **Enabled**
4. Expand the Vertical Layout property
5. Set the Maximum Distance property to **0.25**
6. Set the Alignment property to **Center**
7. Set the Alignment Tolerance property to **0.1**
8. Set the Result Separator property to (a single space)



Step

1. Drag the [Array] PO Number Data Type on top of the Invoice Field Labels Data Type
2. Select the Invoice Field Labels Data Type to ensure the array is now being included in the results

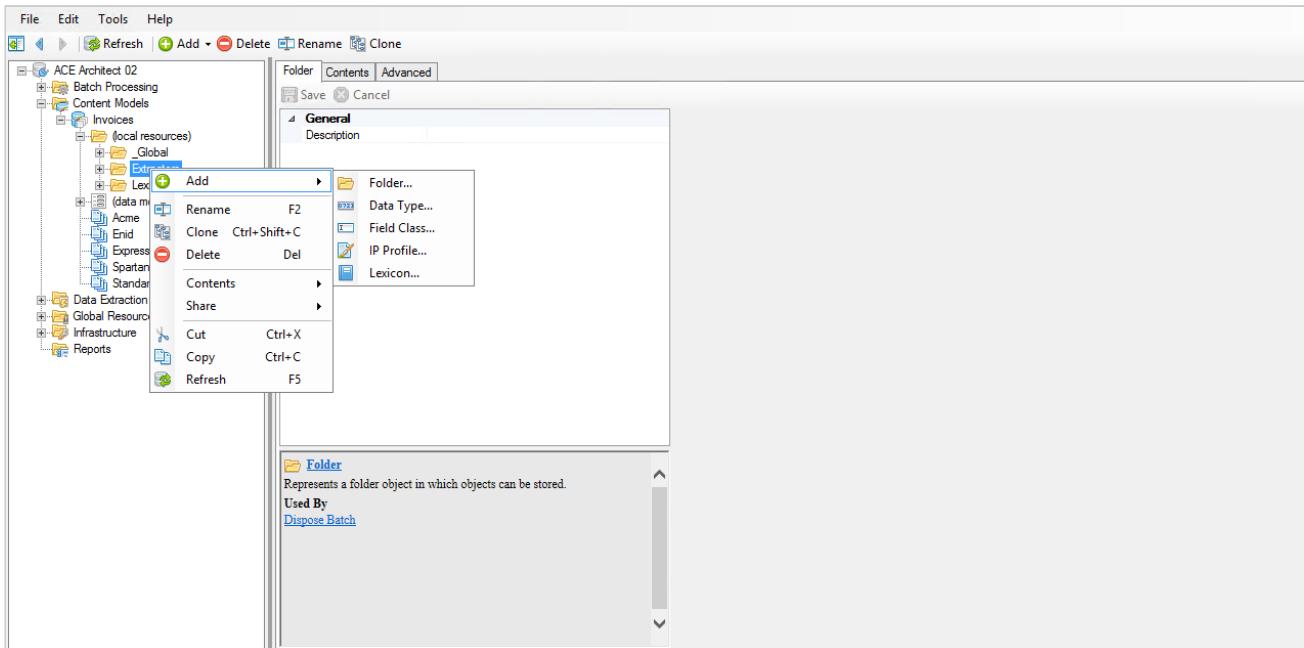


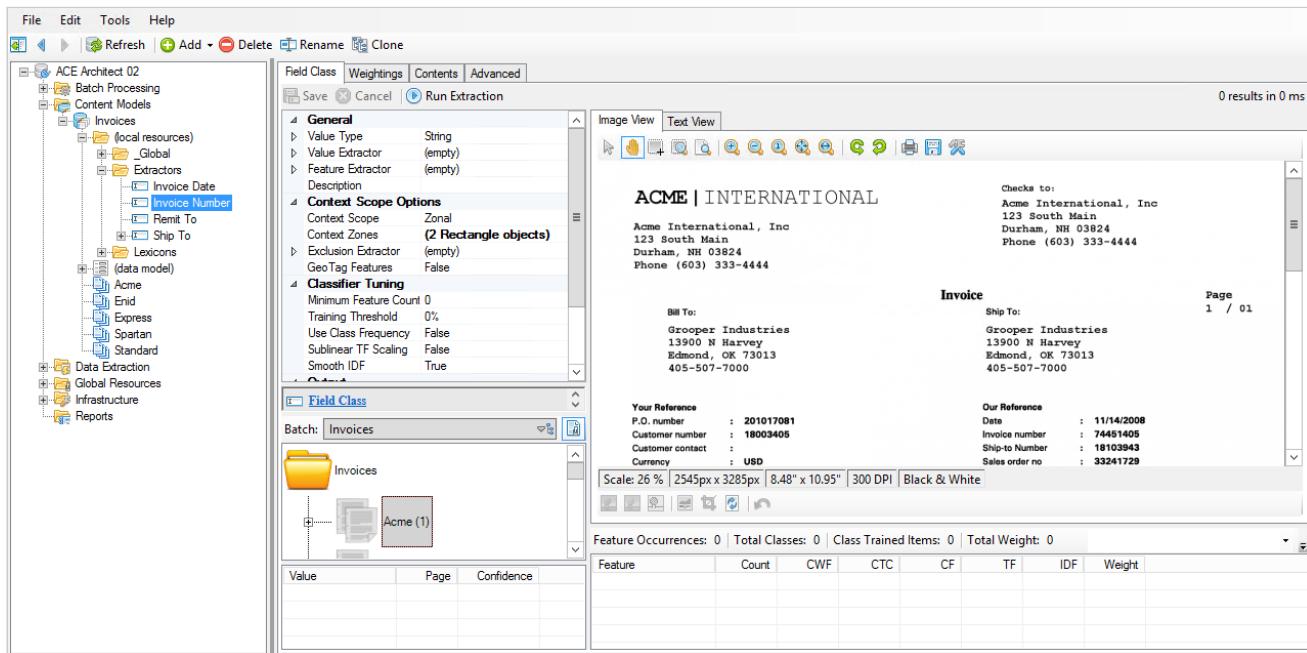


Invoice Number

Step

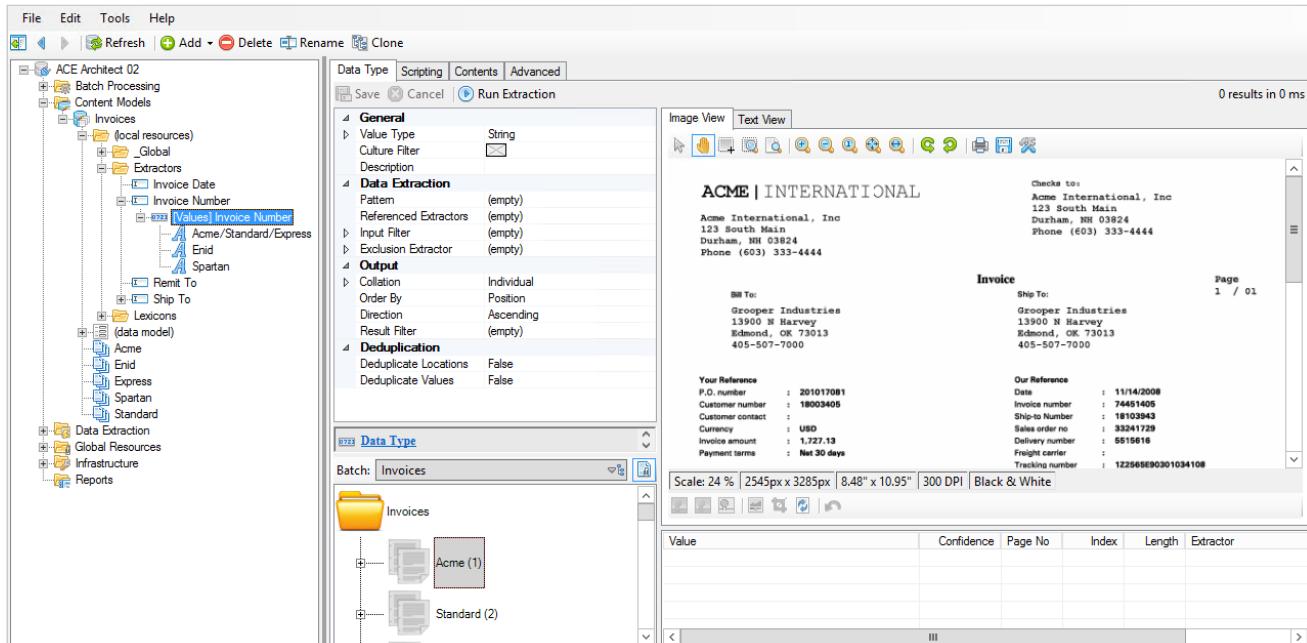
1. Right click on the Extractors folder and select Add > Field Class...
2. Name it **Invoice Number**





Step

1. Create a child Data Type under the Invoice Number Field Class
2. Name it [Values] Invoice Number
3. Create three child Data Formats under the [Values] Invoice Number Data Type and name them:
 - a. Acme/Standard/Express
 - b. Enid
 - c. Spartan



-- Edit the patterns for each of the Data Formats: - Acme/Standard/Express

```
```tab="Value Pattern"
[0-9]{6,16}
```

```

- Enid

```
```tab="Value Pattern"

```

([A-Z]{3})/(@Number){8})

...

```tab="Output Format"

{1}{2:Number}

...

- Spartan

```tab="Value Pattern"

([A-Z]{2})(@Number){2}-(@Number){6})

...

```tab="Output Format"

{1}{2:Number}-{3}

...

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

Batch Processing

Content Models

Invoices

Extractors

Invoice Date

Invoice Number

[Values] Invoice Number

Acme/Standard/Express

Enid

Spartan

Remit To

Ship To

Lexicons (data model)

Acme

Enid

Express

Spartan

Standard

Data Extraction

Global Resources

Infrastructure

Reports

Pattern Editor Properties

Value Pattern

1 [0-9]{6,16}

Look Ahead Pattern

1

Look Behind Pattern

1

Output Format

OK: [0-9]{6,16}

Batch: Invoices

Invoices

Acme (1)

Standard (2)

Express (3)

Image Text

ACME | INTERNATIONAL

Checks to:

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

Invoice

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

Our Reference

P.O. number : 201017081

Customer number : 18003405

Customer contact :

Currency : USD

Invoice amount : 1,727.13

Payment terms : Net 30 days

Date : 11/14/2008

Invoice number : 74451405

Ship-to Number : 18103943

Sales order no : 33241729

Delivery number : 6515616

Freight carrier :

Scale: 25% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

Results (10)

| | Confidence | Page No | Index | Length |
|-----------|------------|---------|-------|--------|
| 201017081 | 100 % | 1 | 406 | |
| 18003405 | 100 % | 1 | 449 | |
| 74451405 | 100 % | 1 | 473 | |
| 18103943 | 100 % | 1 | 515 | |

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

Batch Processing

Content Models

Invoices

Extractors

Invoice Date

Invoice Number

[Values] Invoice Number

Acme/Standard/Express

Enid

Spartan

Remit To

Ship To

Lexicons (data model)

Acme

Enid

Express

Spartan

Standard

Data Extraction

Global Resources

Infrastructure

Reports

Pattern Editor Properties

Value Pattern

1 ([A-Z]{3})/(@Number){8}

Look Ahead Pattern

1

Look Behind Pattern

1

Output Format

{1}{2:Number}

OK: ([A-Z]{3})/([0-9]{8})

Batch: Invoices

Express (3)

Spartan (4)

Enid (5)

Express (6)

Image Text

Enid Parts

INVOICE

INVOICE DATE: 12-05-08 INVOICE #: CS1/10076138 PAGE: 1

PH: (405) 555-1212 FAX: (405) 444-2121

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

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

SHIP: Grooper Industries
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 |
|--------------|------------|-----------------|--------------|-----------------------|----------|-----------|--------|
| 100049766 | 11-11-08 | CUS001410 | 099325 | 010013704 | FXG | 12-03-08 | Net 30 |

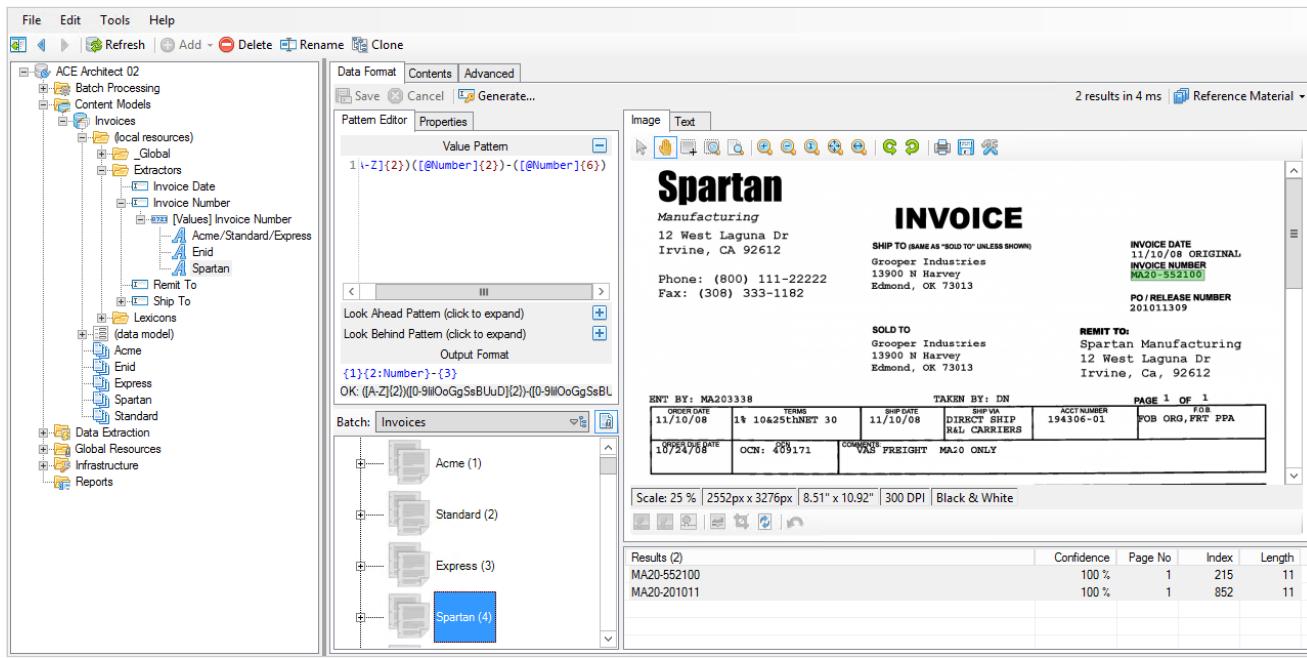
ITEM NUMBER ITEM DESCRIPTION

QUANTITY ORDERED QUANTITY SHIPPED STK UNIT PRICE UNIT DISCOUNT EXTENDED PRICE

Scale: 25% | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

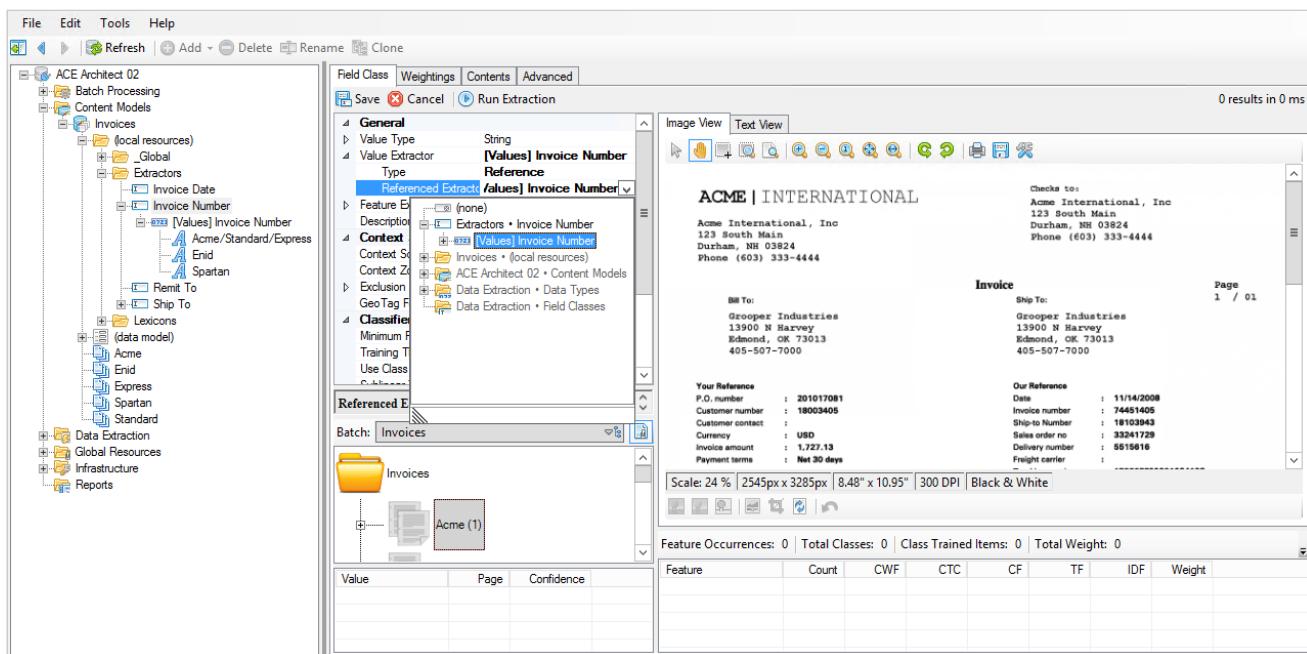
Results (1)

| | Confidence | Page No | Index | Length |
|--------------|------------|---------|-------|--------|
| CS1/10076138 | 100 % | 1 | 49 | 12 |



Step

1. Select the Invoice Number Field Class
2. In the General section, set the Value Extractor > Type property to Reference
3. Set the Referenced Extractor property to the Extractors • Invoice Number > [Values] Invoice Number Data Type
4. Set the Feature Extractor > Type property to Reference
5. Set the Referenced Extractor property to the Invoices • (local resources) > _Global > Invoice Field Labels Data Type
6. Select Acme (1)
7. Click Run Extraction



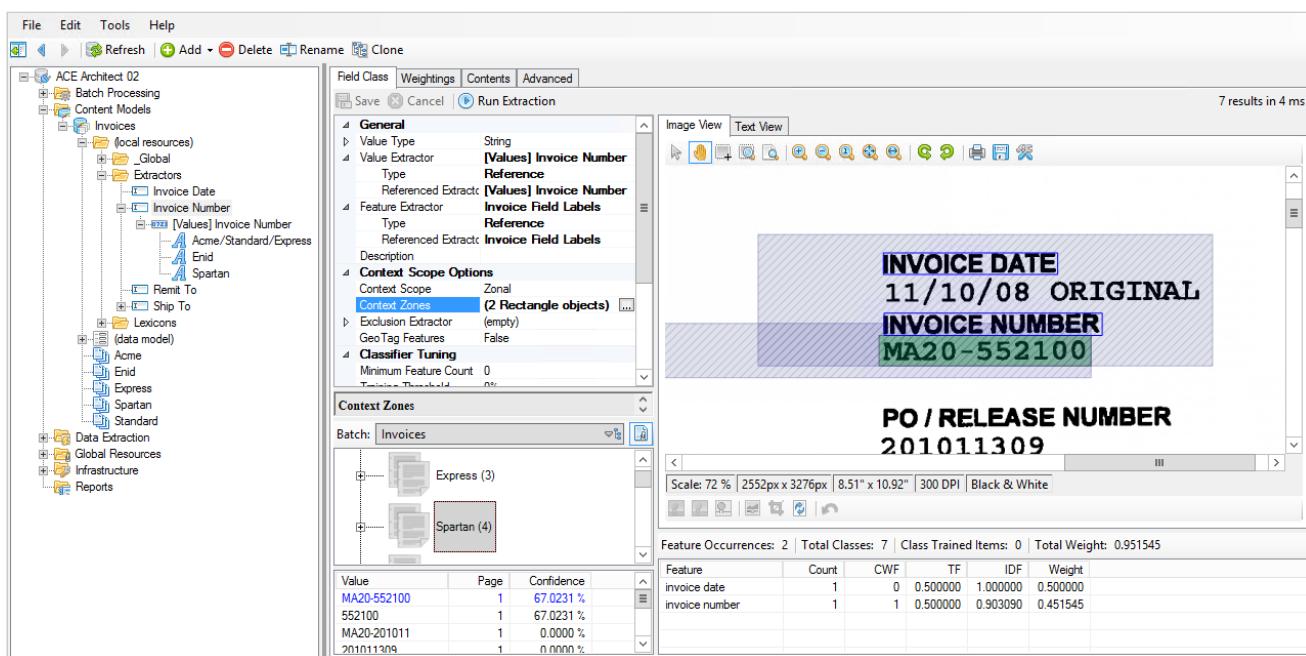
Step

1. In the candidate list, find the value that has the appropriate feature(s) to train

2. Right click and select Train As Positive

Step

1. Select Spartan (4)
2. Click to edit the Context Zones property
3. In the Context Zones window, edit the dimensions as follows:
 - a. Zone 1:
 - i. Left: -4
 - ii. Top: -0.02
 - iii. Right: 0.35
 - iv. Bottom: 0.1
 - b. Zone 2:
 - i. Left: -1
 - ii. Top: -0.25
 - iii. Right: 0.375
 - iv. Bottom: 0.05
4. Click OK



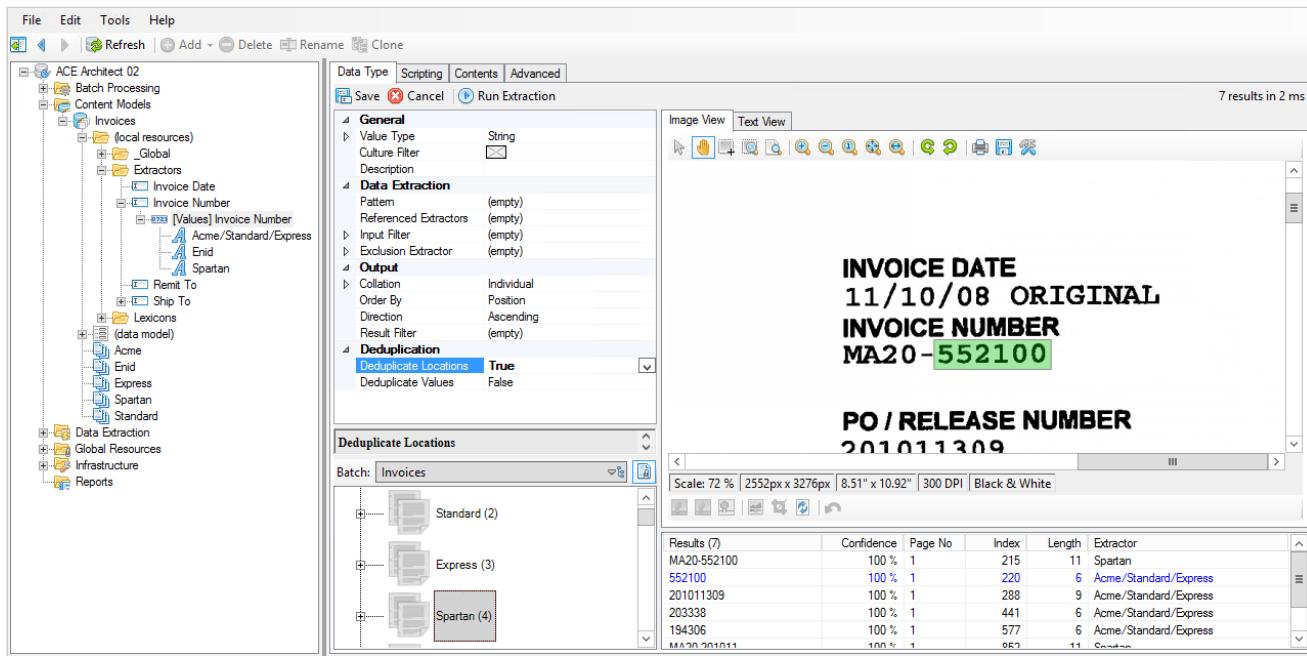
Screenshot of ACE Architect 02 interface showing the extraction configuration for an invoice number. The 'Field Class' tab is selected, showing a 'General' section with 'Value Type' set to 'String' and 'Value Extractor' set to '[Values] Invoice Number'. The 'Image View' tab shows a snippet of an invoice document with the text 'ORIGINAL MEMBER 100'. The 'Context Zones' dialog is open, displaying two zones: Zone 1 at (-4, -0.02)(0.35, 0.1) and Zone 2 at (-1, -0.5)(1.0). The 'Feature Occurrences' table shows two occurrences of 'invoice date' and one occurrence of 'invoice number'.

Screenshot of ACE Architect 02 interface showing the extraction configuration for an invoice number. The 'Field Class' tab is selected, showing a 'General' section with 'Value Type' set to 'String' and 'Value Extractor' set to '[Values] Invoice Number'. The 'Image View' tab shows a snippet of an invoice document with the text 'ORIGINAL MEMBER 100'. The 'Context Zones' dialog is open, displaying two zones: Zone 1 at (-4, -0.02)(0.35, 0.1) and Zone 2 at (-1, -0.25)(0.375, 0.05). The 'Feature Occurrences' table shows two occurrences of 'invoice date' and one occurrence of 'invoice number'.

Screenshot of ACE Architect 02 interface showing the extraction configuration for an invoice number. The 'Field Class' tab is selected, showing a 'General' section with 'Value Type' set to 'String' and 'Value Extractor' set to '[Values] Invoice Number'. The 'Image View' tab shows a snippet of an invoice document with the text 'INVOICE DATE 11/10/08 ORIGINAL INVOICE NUMBER MA20-552100'. The 'Context Zones' dialog is open, showing 'Context Scope Options' with 'Context Scope' set to 'Zonal' and 'Context Zones' set to '(2 Rectangle objects)'. The 'Feature Occurrences' table shows one occurrence of 'invoice number'.

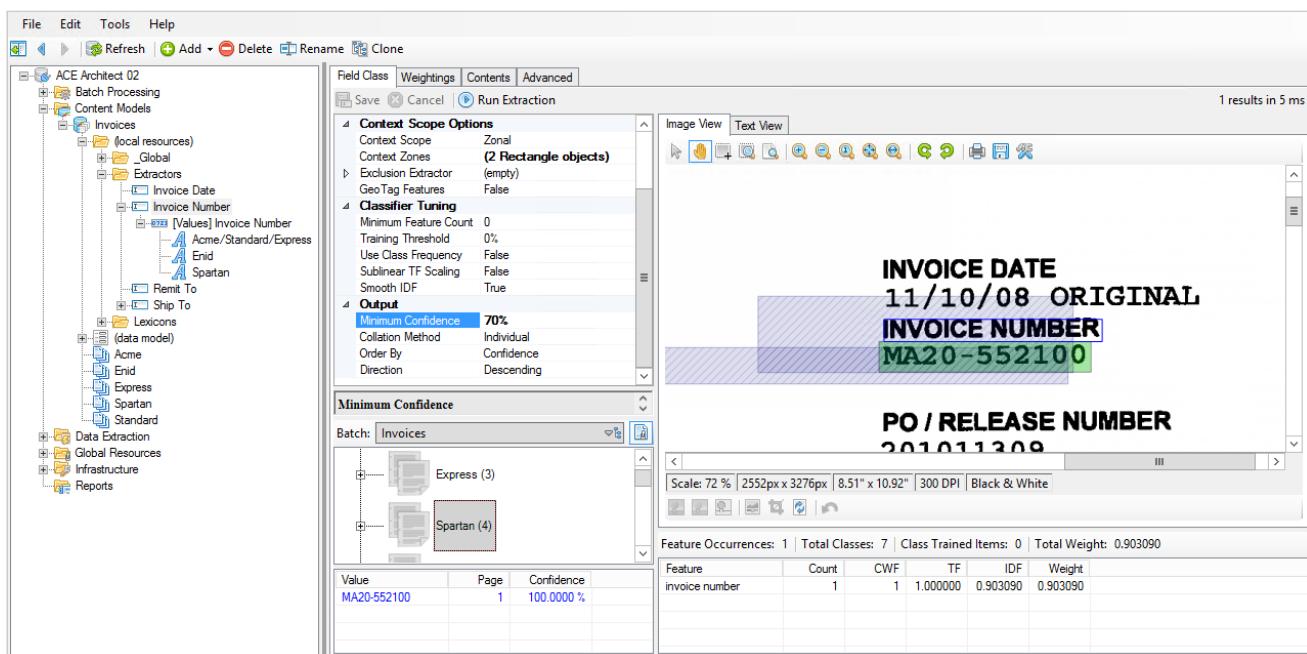
Step

- Click on the [Values] Invoice Number Data Type
- In the Deduplication section, set the Deduplicate Locations property to **True**



Step

1. Click on the Invoice Number Field Class
2. In the Output section, set the Minimum Confidence property to 70%



Step

1. In the Data Model, select the Invoice Number Data Field
2. In the General section, set the Value Extractor > Type property to Reference
3. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > Invoice Number Field Class.
4. In the Behavior section, set the Required property to True

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
 - Acme
 - End
 - Express
 - Spartan
 - Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **Cancel** **Test Extraction**

General

- Value Type: String
- Value Extractor: (empty)
- Minimum Confidence: 20%
- Sub-Element Name: Description
- Appearance

 - Alignment: Left
 - Display Width: 100
 - Error Color: 255, 192, 192
 - Background Color: Window
 - Foreground Color:WindowText
 - Visible: True
 - Label Position: Left

Data Field

Represents a field.

Remarks

Fields can be children of [Data Model](#) or [Data Section](#) objects. Field objects define characteristics such as:

- The appearance of the item in the [Index Panel](#).
- The default ESP™ Extraction settings. These settings can be overridden on

Batch: Invoices

Document View **Text View** **Instance View**

Invoice Number

ACME | INTERNATIONAL

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

Invoice

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

Page 1 / 01

Scale: 19% **2545px x 3285px** **8.48" x 10.95"** **300 DPI** **Black & White**

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
 - Acme
 - End
 - Express
 - Spartan
 - Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **X Cancel** **Test Extraction**

General

- Value Type: String
- Value Extractor: **Referenced Extractor**
- Minimum Confidence: 20%
- Sub-Element Name: Description
- Appearance

 - Alignment: Left
 - Display Width: 100
 - Error Color: 255, 192, 192
 - Background Color: Window
 - Foreground Color:WindowText
 - Visible: True
 - Label Position: Left

Referenced Extractor

Type: [Grooper Node](#)

When Type is set to 'Reference', a reference to another object is created.

Property Type: [Grooper Node](#)

Grooper Nodes are objects which at every level of the tree structure. All Grooper Node objects support a base set of properties and behaviors common to

Batch: Invoices

Document View **Text View** **Instance View**

Invoice Number

ACME | INTERNATIONAL

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

Invoice

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

Page 1 / 01

Scale: 19% **2545px x 3285px** **8.48" x 10.95"** **300 DPI** **Black & White**

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
 - Acme
 - End
 - Express
 - Spartan
 - Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **X Cancel** **Test Extraction**

Label Position: Left **Label Alignment:** MiddleLeft

Behavior

- FieldSense™: Disabled
- Input Mask: Disabled
- Multi Line: Disabled
- Tooltip: (Text)
- Read Only: False
- Required:** **True**
- Sticky: False
- Requires Validation: False
- Character Casing: Normal
- Preserve Character Data: False

Required

Type: [Boolean](#). Default: [False](#)

If set to true, a value is required in the field.

Batch: Invoices

Document View **Text View** **Instance View**

Invoice Number

ACME | INTERNATIONAL

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

Invoice

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

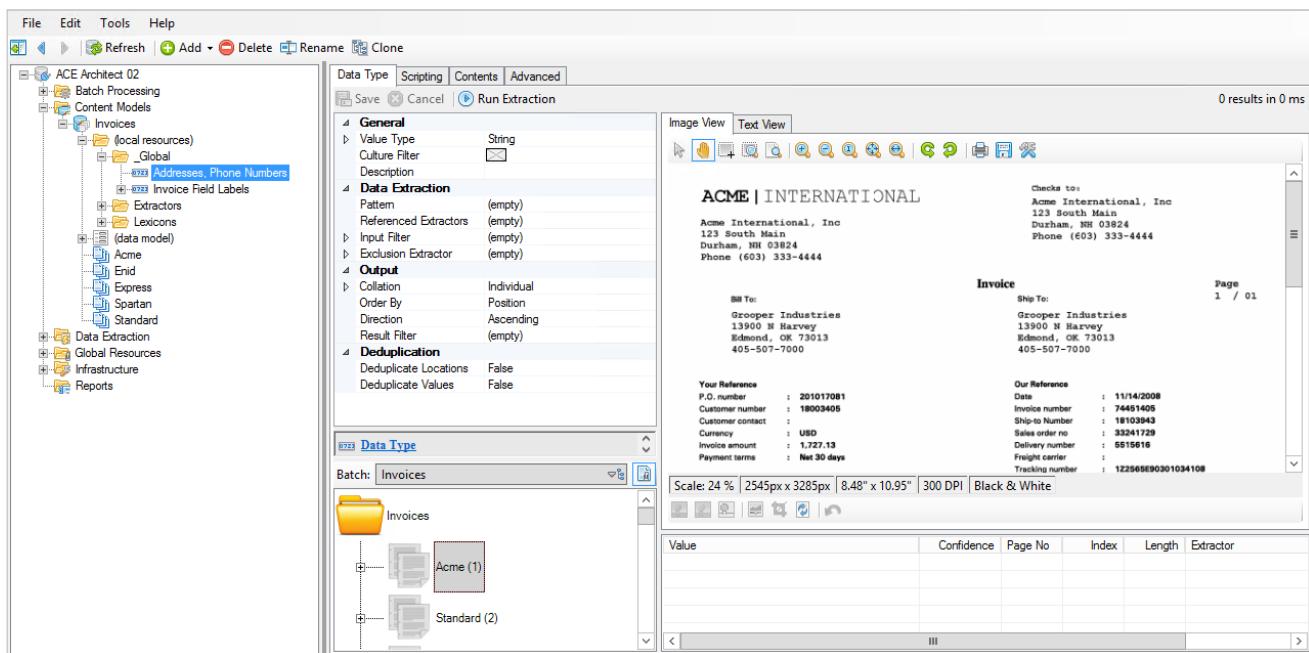
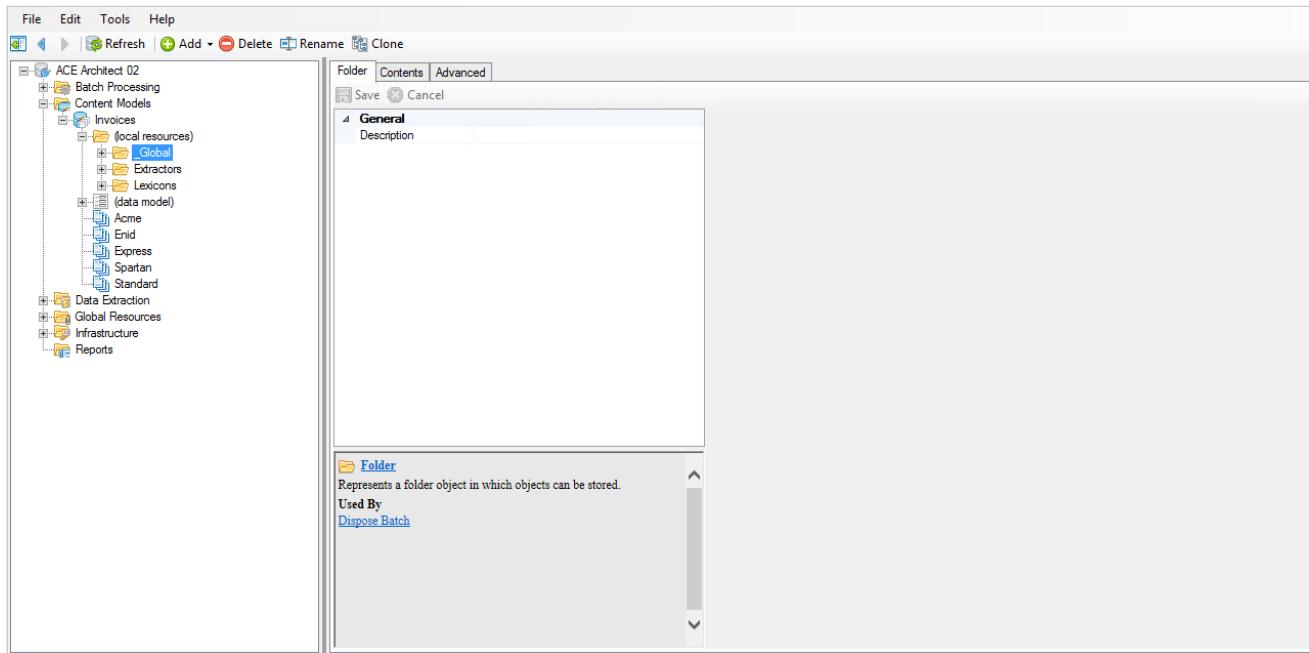
Page 1 / 01

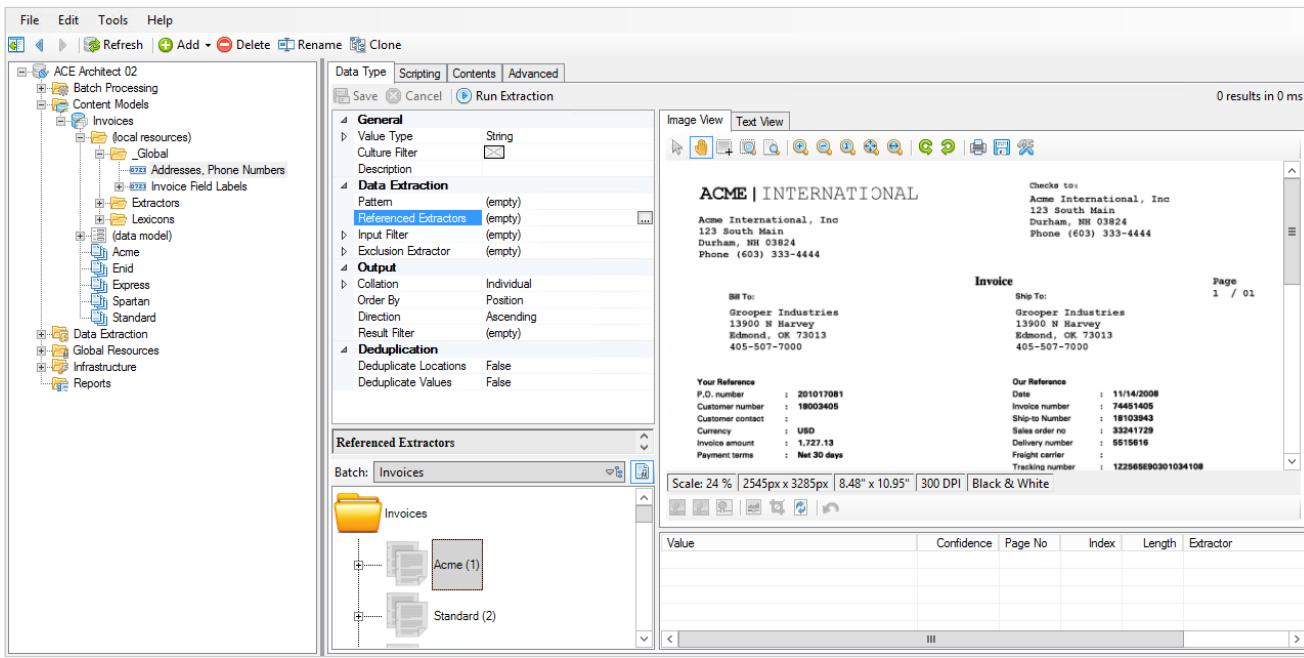
Scale: 19% **2545px x 3285px** **8.48" x 10.95"** **300 DPI** **Black & White**

PO Number/Exclusion Extractor

Step

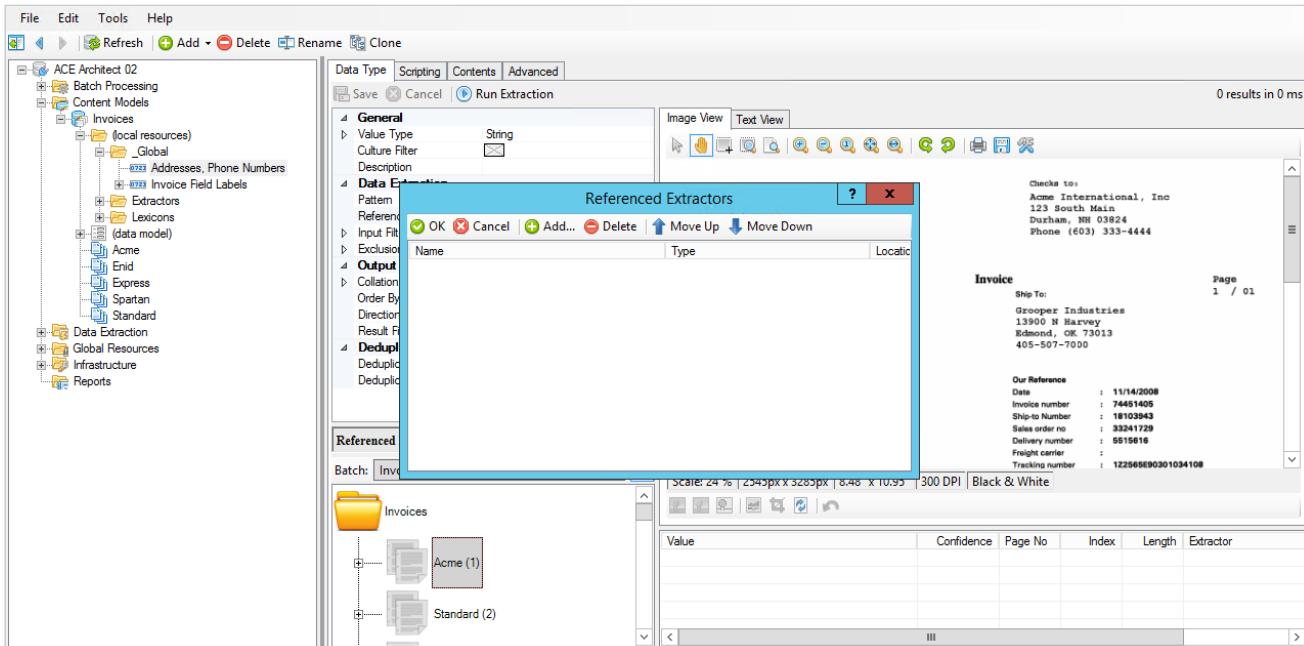
1. Right click on the (local resources) > _Global folder and select Add > Data Type...
2. Name it Addresses, Phone Numbers
3. In the Data Extraction section, click to edit Referenced Extractors

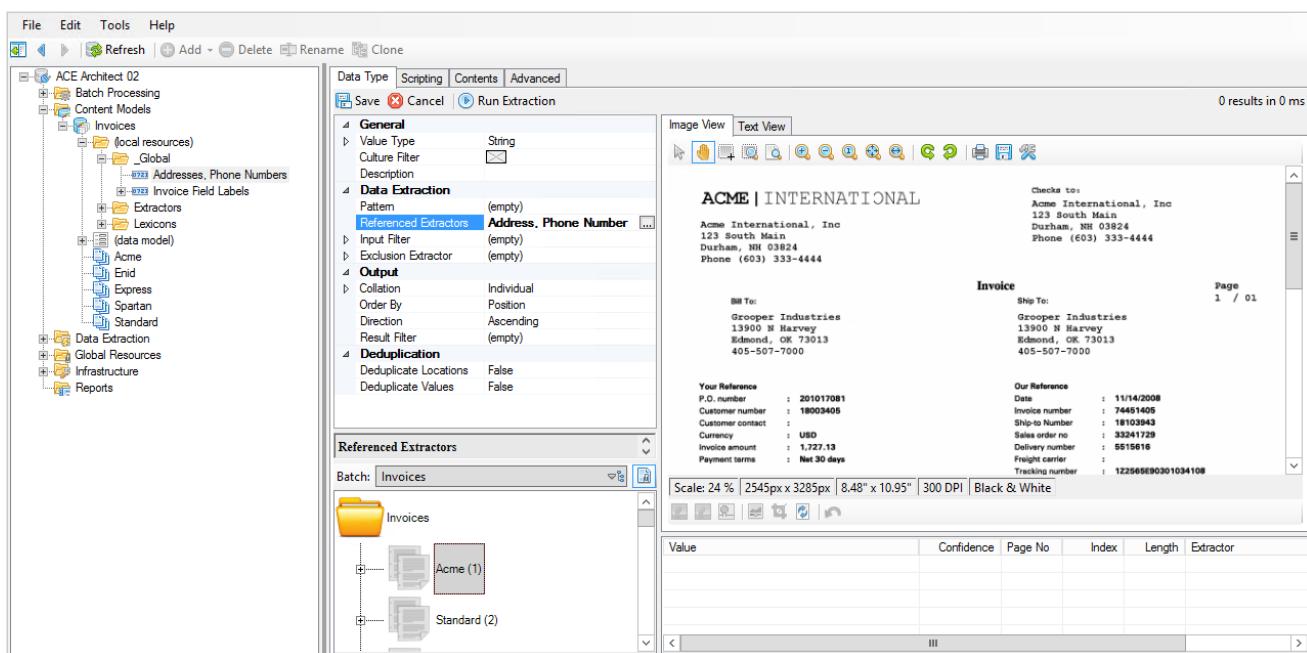
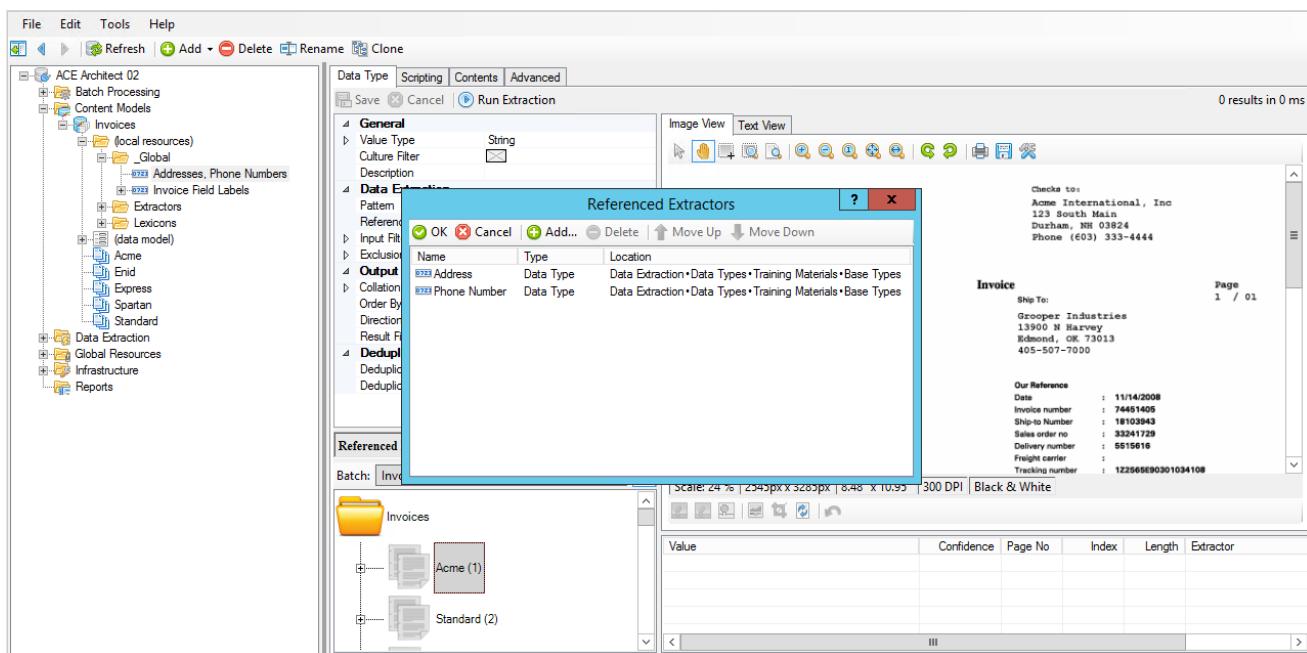
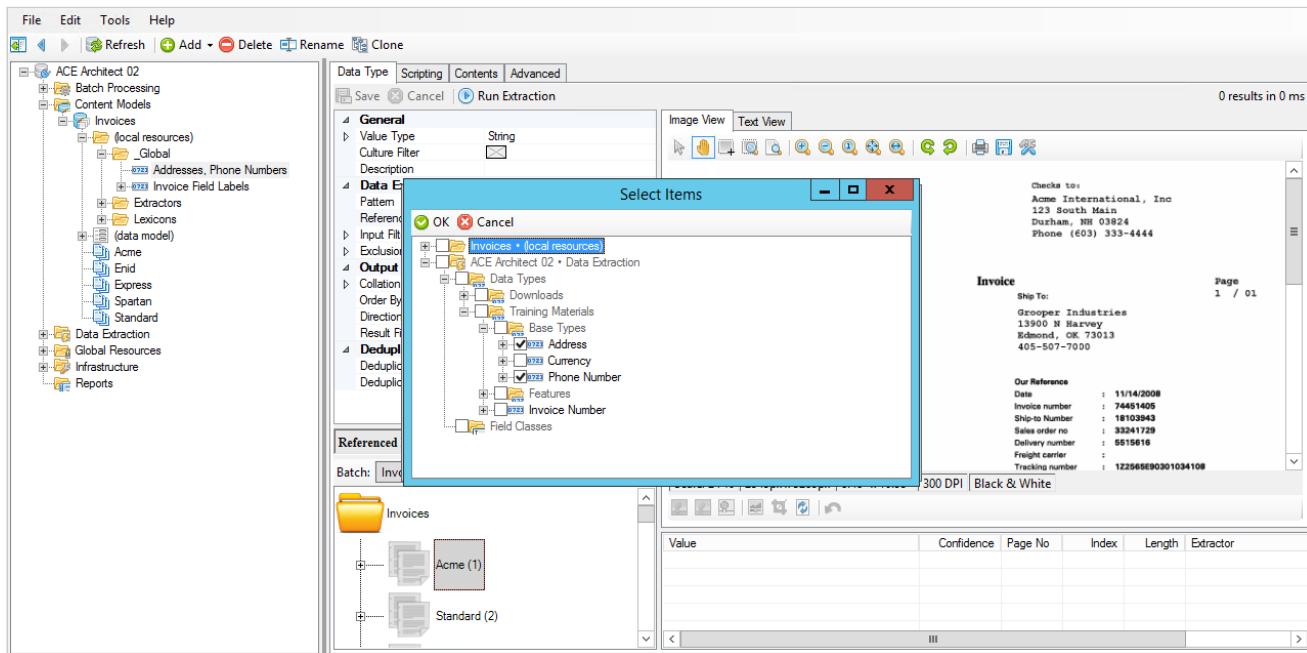




Step

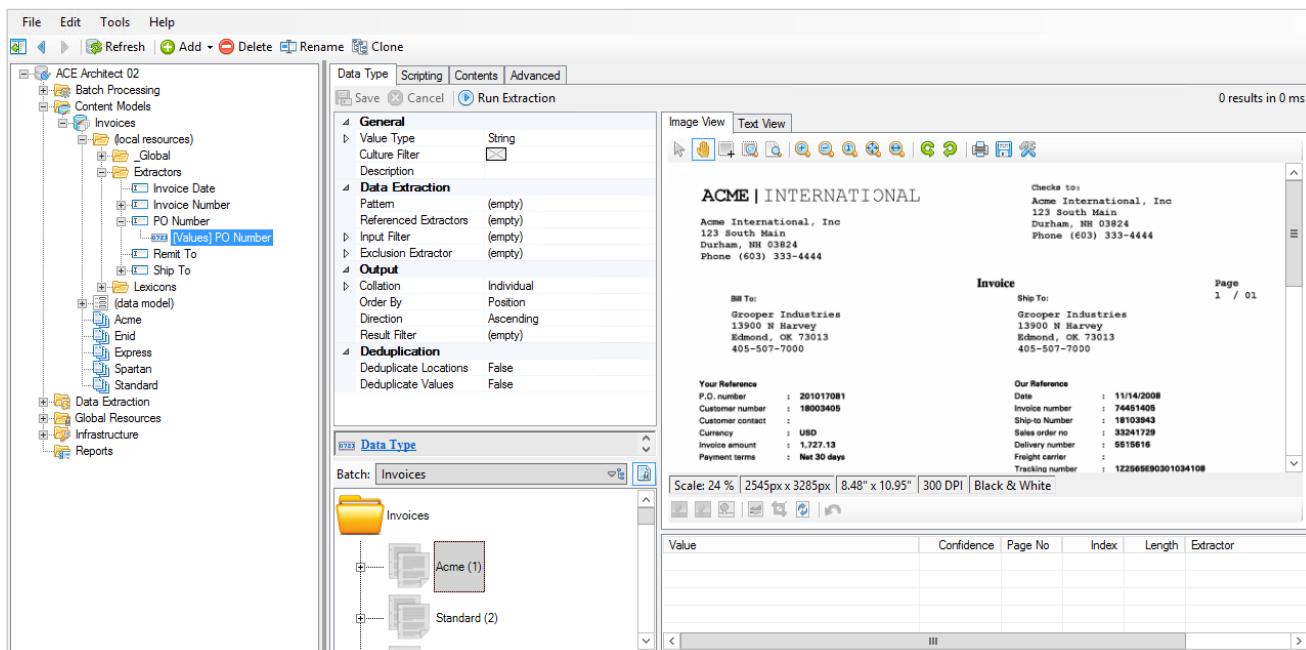
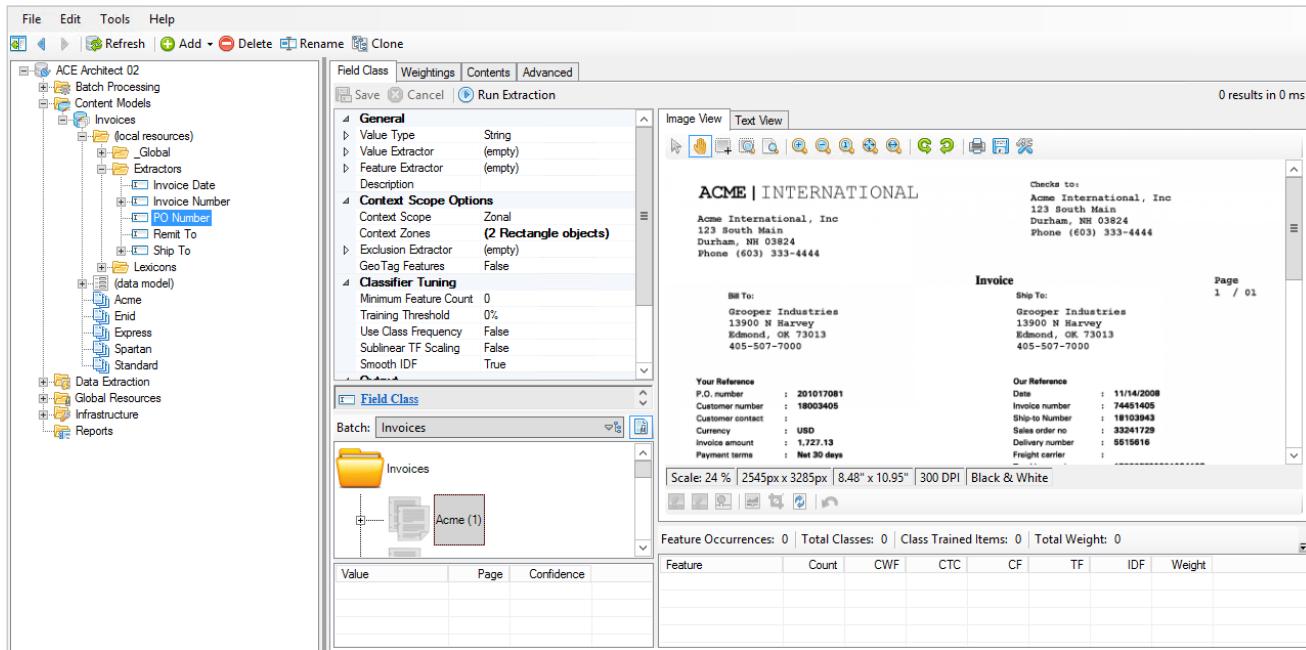
1. In the Referenced Extractors window, click Add...
2. In the Select Items window, expand (root) • Data Extraction > Data Types > Training Materials > Base Types and check the following:
 - a. Address
 - b. Phone Number
3. Click OK to close the Select Items window
4. Click OK to close the Referenced Extractors window

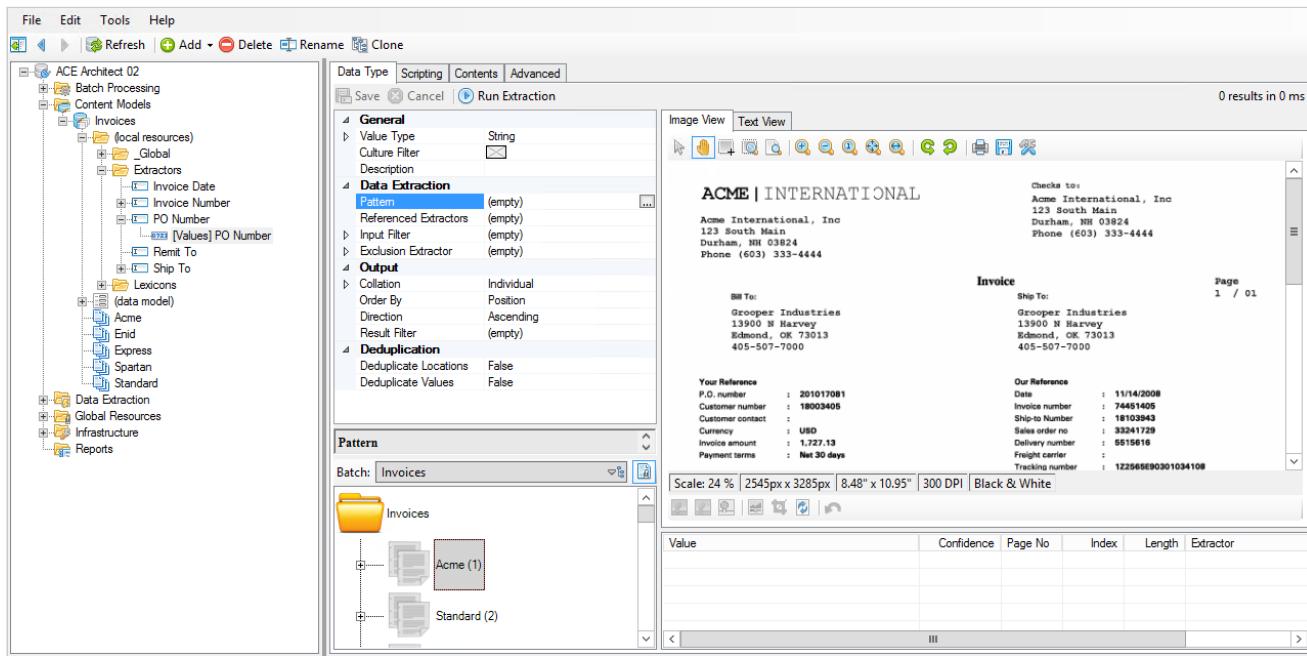




Step

1. Right click on the Extractors folder and select Add > Field Class...
2. Name it **PO Number**
3. Create a child Data Type of PO Number and name it **[Values] PO Number**
4. On the **[Values] PO Number** Data Type, click to edit the Pattern



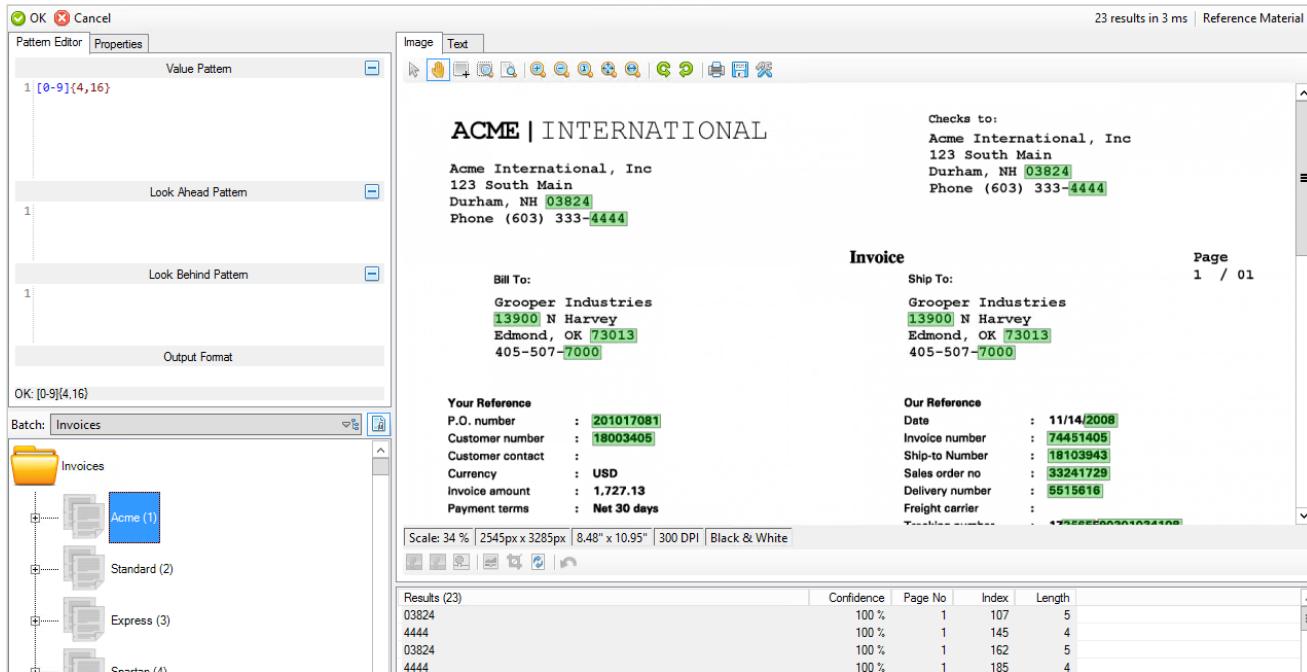


Step

1. In the Value Pattern editor, type:

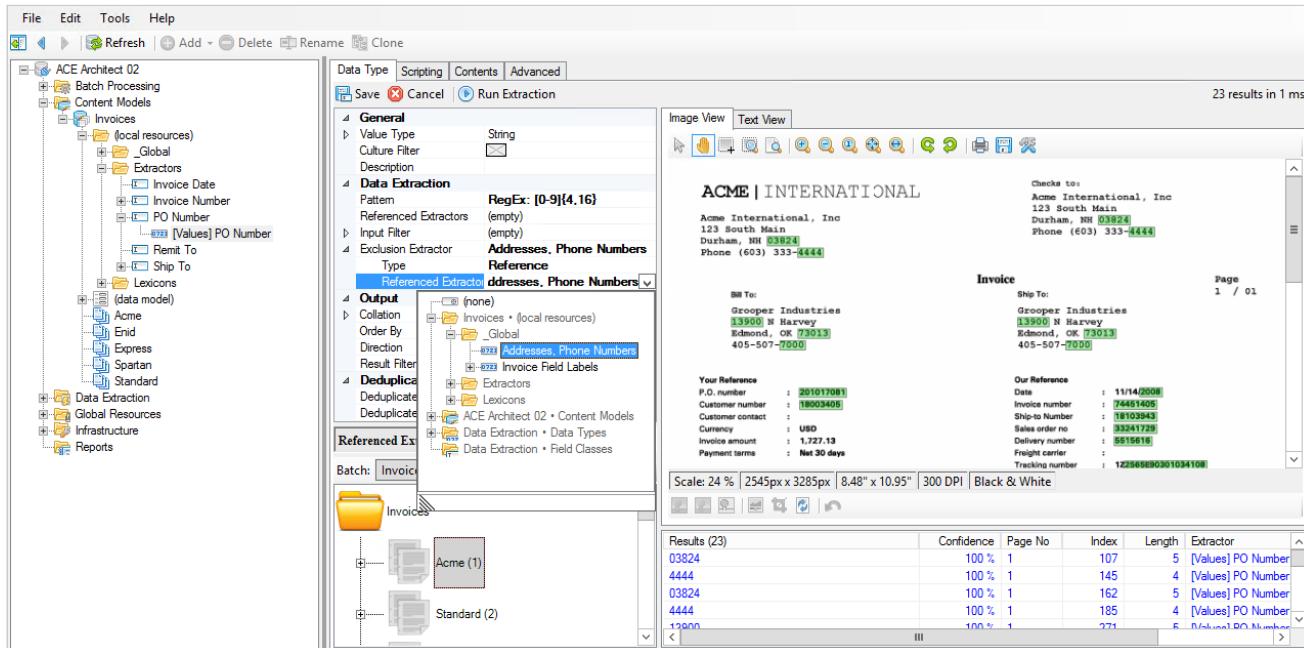
[0-9]{4,16}

2. Click OK



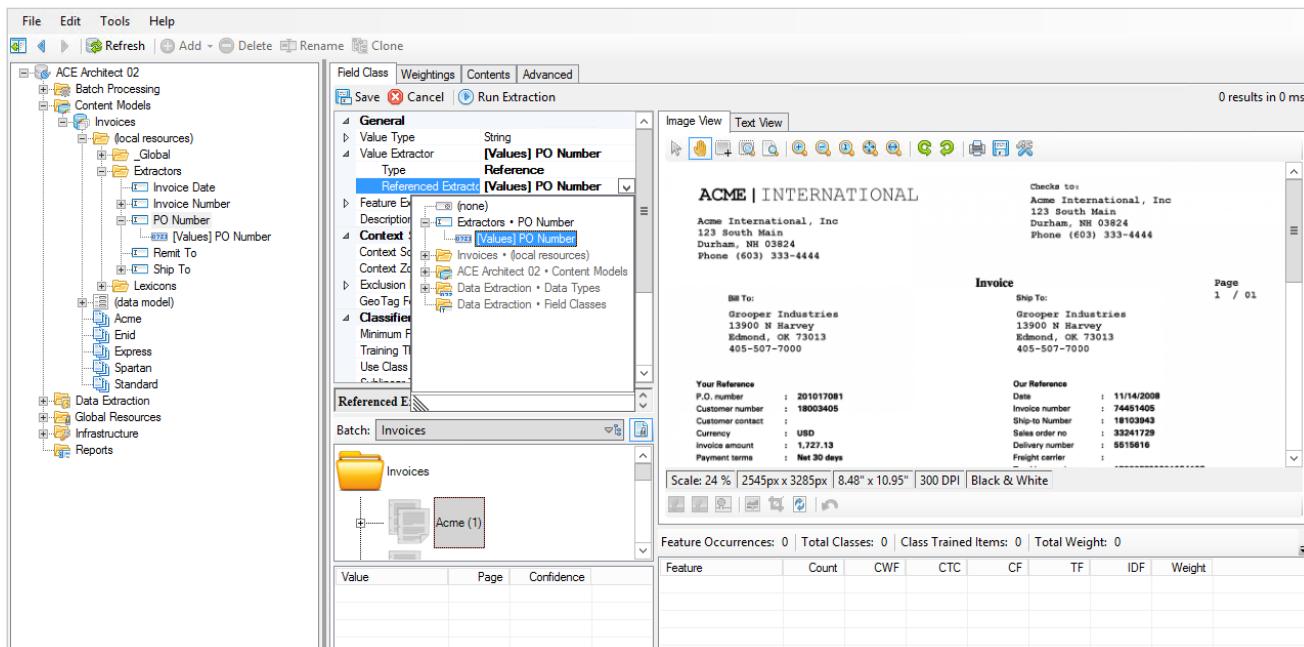
Step

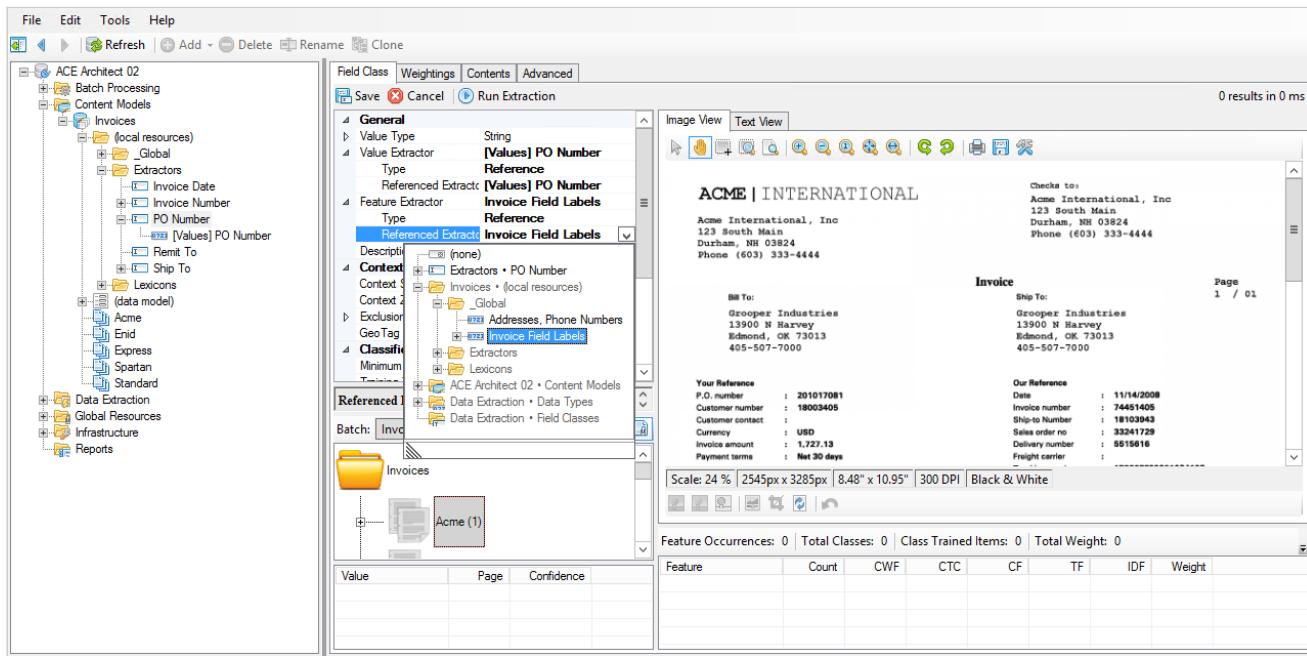
1. On the [Values] PO Number Data Type, in the Data Extraction section, set the Exclusion Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Invoices • (local resources) > _Global > Addresses, Phone Numbers Data Type



> Step

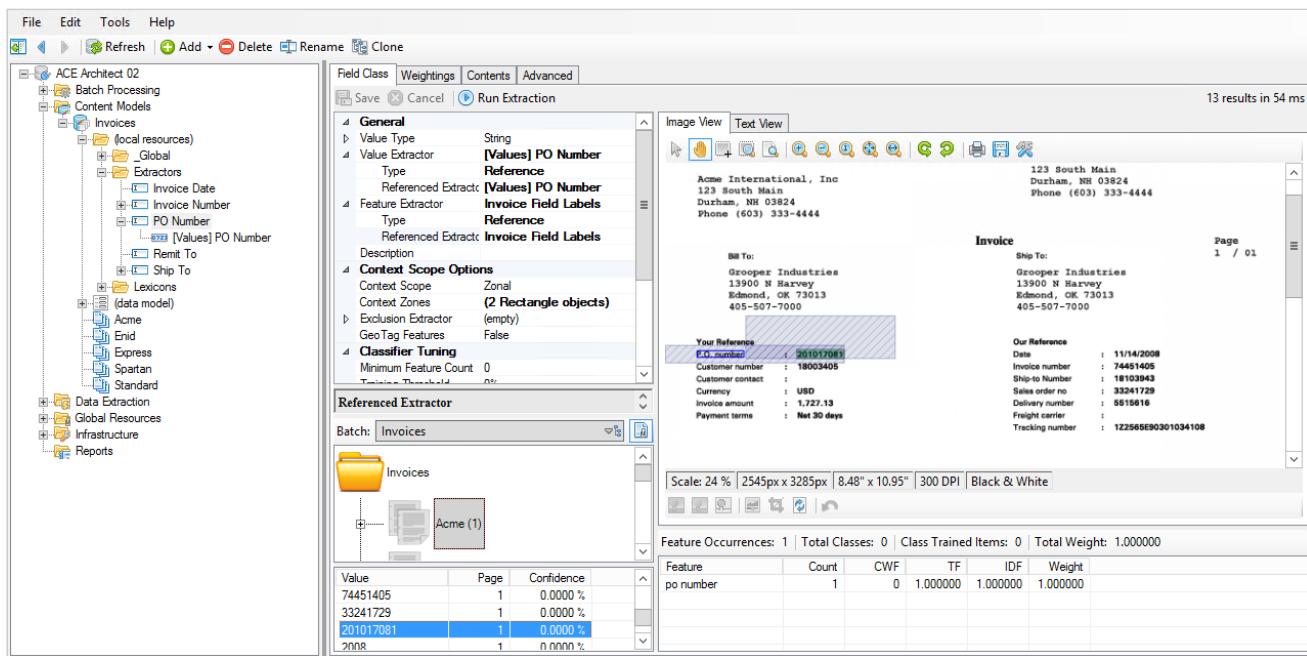
1. Click on the PO Number Field Class
 2. In the General section, set the Value Extractor > Type property to Reference
 3. Set the Referenced Extractor property to the Extractors • PO Number > **[Values] PO Number** Data Type
 4. In the General section, set the Feature Extractor > Type property to Reference
 5. Set the Referenced Extractor property to the Invoices • (local resources) > Global > Invoice Field Labels Data Type

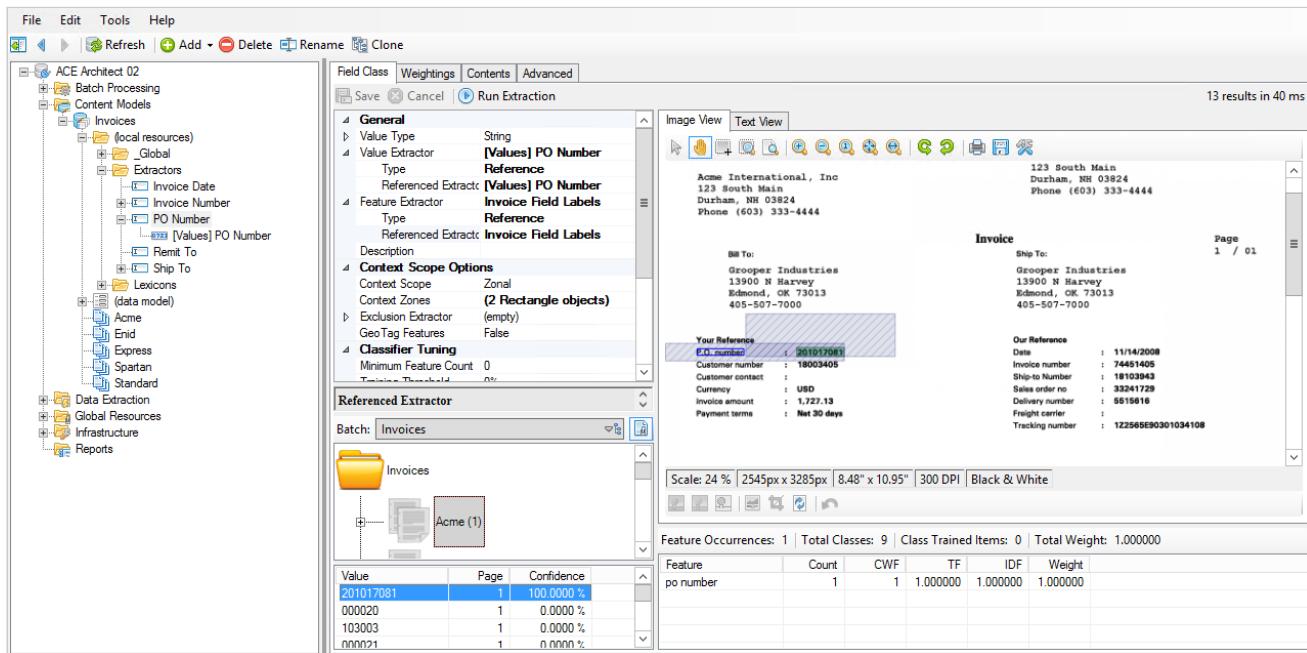




Step

1. Select Acme (1)
2. Click Run Extraction
3. In the candidate list, find the value that has the appropriate feature(s) to train
4. Right click and select Train As Positive





Step

1. Click to edit the Context Zones property
2. In the Context Zones window, edit the dimensions as follows:
 - a. Zone 1:
 - i. Left: -3.278
 - ii. Top: -0.054
 - iii. Right: 0.312
 - iv. Bottom: 0.054
 - b. Zone 2:
 - i. Left: -0.67
 - ii. Top: -0.37
 - iii. Right: 0.5
 - iv. Bottom: 0.04
3. Click OK
4. In the Output section, set the Minimum Confidence property to 70%

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - _Global
 - Extractors
 - Invoice Date
 - Invoice Number
 - PO Number
 - [Values] PO Number
 - Remit To
 - Ship To
 - Lexicons (data model)
 - Acme
 - End
 - Express
 - Spartan
 - Standard
 - Data Extraction
 - Global Resources
 - Infrastructure
 - Reports

Field Class Weightings Contents Advanced

Save Cancel Run Extraction

16 results in 33 ms

General

 - Value Type String
 - Value Extractor [Values] PO Number Reference Referenced Extract: [Values] PO Number
 - Feature Extractor Invoice Field Labels Type Reference Referenced Extract: Invoice Field Labels Description
 - Context Scope Options Context Scope Zonal
 - Context Zones (2 Rectangle objects) ...
 - Exclusion Extractor (empty)
 - Geo-Tag Features False
 - Classifier Tuning Minimum Feature Count 0

Context Zones

Batch: Invoices

Acme (1)
Standard (2)

| Value | Page | Confidence |
|-------------|------|------------|
| 209001161 | 1 | 57.7350 % |
| 2008 | 1 | 0.0000 % |
| 3018 | 1 | 0.0000 % |
| 7032824214n | 1 | 0.0000 % |

Image View Text View

ACCOUNT NUMBER
INVOICE NUMBER
INVOICE DATE
DUE DATE
AMOUNT DUE

PO NUMBER: 209001161

Scale: 73 % 2545px x 3274px 8.48" x 10.91" 300 DPI Black & White

Feature Occurrences: 3 | Total Classes: 9 | Class Trained Items: 0 | Total Weight: 1.000000

| Feature | Count | CWF | TF | IDF | Weight |
|----------------|-------|-----|----------|----------|----------|
| po number | 1 | 1 | 0.333333 | 1.000000 | 0.333333 |
| invoice amount | 1 | 0 | 0.333333 | 1.000000 | 0.333333 |
| due date | 1 | 0 | 0.333333 | 1.000000 | 0.333333 |

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - _Global
 - Extractors
 - Invoice Date
 - Invoice Number
 - PO Number
 - [Values] PO Number
 - Remit To
 - Ship To
 - Lexicons (data model)
 - Acme
 - End
 - Express
 - Spartan
 - Standard
 - Data Extraction
 - Global Resources
 - Infrastructure
 - Reports

Field Class Weightings Contents Advanced

Save Cancel Run Extraction

16 results in 33 ms

General

 - Value Type String
 - Value Extractor [Values] PO Number Reference Referenced Extract: [Values] PO Number
 - Feature Extractor Invoice Field Labels Type Reference Referenced Extract: Invoice Field Labels Description
 - Context Scope Options Context Scope Zonal
 - Context Zones (2 Rectangle objects) ...
 - Exclusion Extractor (empty)
 - Geo-Tag Features False
 - Classifier Tuning Minimum Feature Count 0

Context Zones

Zone No Region

Zone 1 (-3.278,-0.054)(0.312,0.054)
Zone 2 (-1,-0.5)(1,0)

| Left | Top | Right | Bottom | Width | Height |
|--------|--------|-------|--------|-------|--------|
| -3.278 | -0.054 | 0.312 | 0.054 | 3.59 | 0.108 |

Left
Type: Double Default: 0
The left edge of the rectangle.

Feature Occurrences: 3 | Total Classes: 9 | Class Trained Items: 0 | Total Weight: 1.000000

| Feature | Count | CWF | TF | IDF | Weight |
|----------------|-------|-----|----------|----------|----------|
| po number | 1 | 1 | 0.333333 | 1.000000 | 0.333333 |
| invoice amount | 1 | 0 | 0.333333 | 1.000000 | 0.333333 |
| due date | 1 | 0 | 0.333333 | 1.000000 | 0.333333 |

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - _Global
 - Extractors
 - Invoice Date
 - Invoice Number
 - PO Number
 - [Values] PO Number
 - Remit To
 - Ship To
 - Lexicons (data model)
 - Acme
 - End
 - Express
 - Spartan
 - Standard
 - Data Extraction
 - Global Resources
 - Infrastructure
 - Reports

Field Class Weightings Contents Advanced

Save Cancel Run Extraction

16 results in 33 ms

General

 - Value Type String
 - Value Extractor [Values] PO Number Reference Referenced Extract: [Values] PO Number
 - Feature Extractor Invoice Field Labels Type Reference Referenced Extract: Invoice Field Labels Description
 - Context Scope Options Context Scope Zonal
 - Context Zones (2 Rectangle objects) ...
 - Exclusion Extractor (empty)
 - Geo-Tag Features False
 - Classifier Tuning Minimum Feature Count 0

Context Zones

Zone No Region

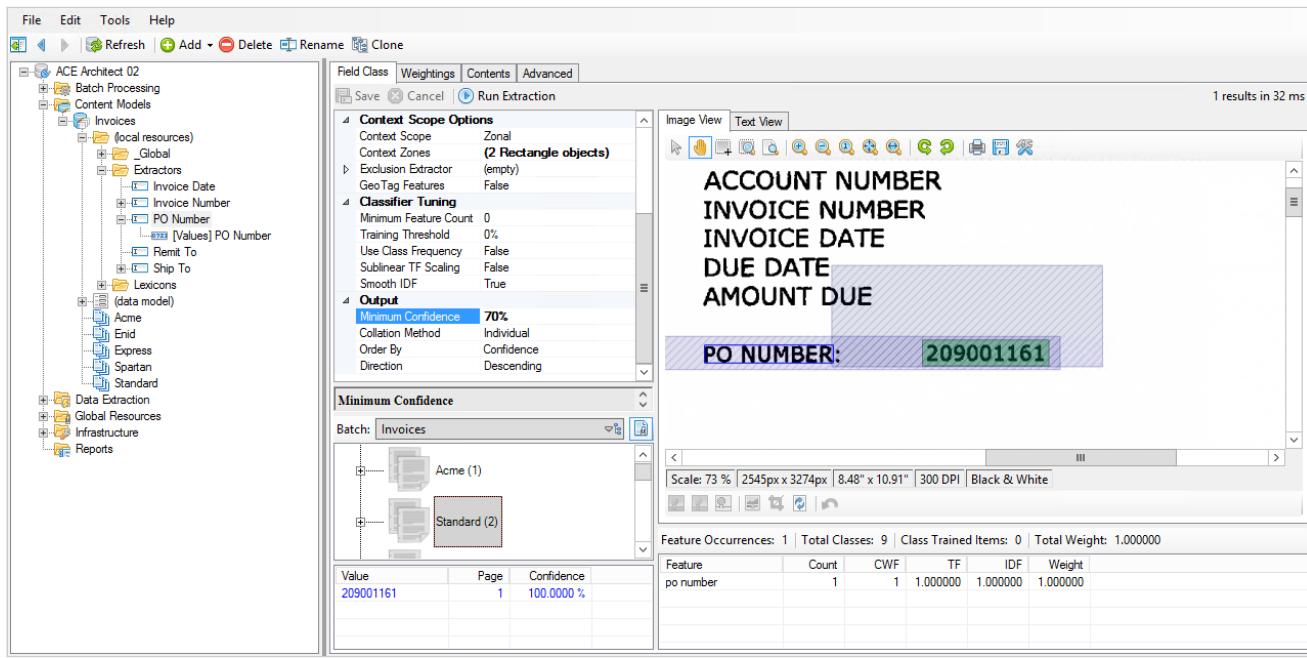
Zone 1 (-3.278,-0.054)(0.312,0.054)
Zone 2 (-0.67,-0.37)(0.5,0.04)

| Left | Top | Right | Bottom | Width | Height |
|--------|--------|-------|--------|-------|--------|
| -3.278 | -0.054 | 0.312 | 0.054 | 3.59 | 0.108 |
| -0.67 | -0.37 | 0.5 | 0.04 | 1.17 | 0.41 |

Left
Type: Double Default: 0
The left edge of the rectangle.

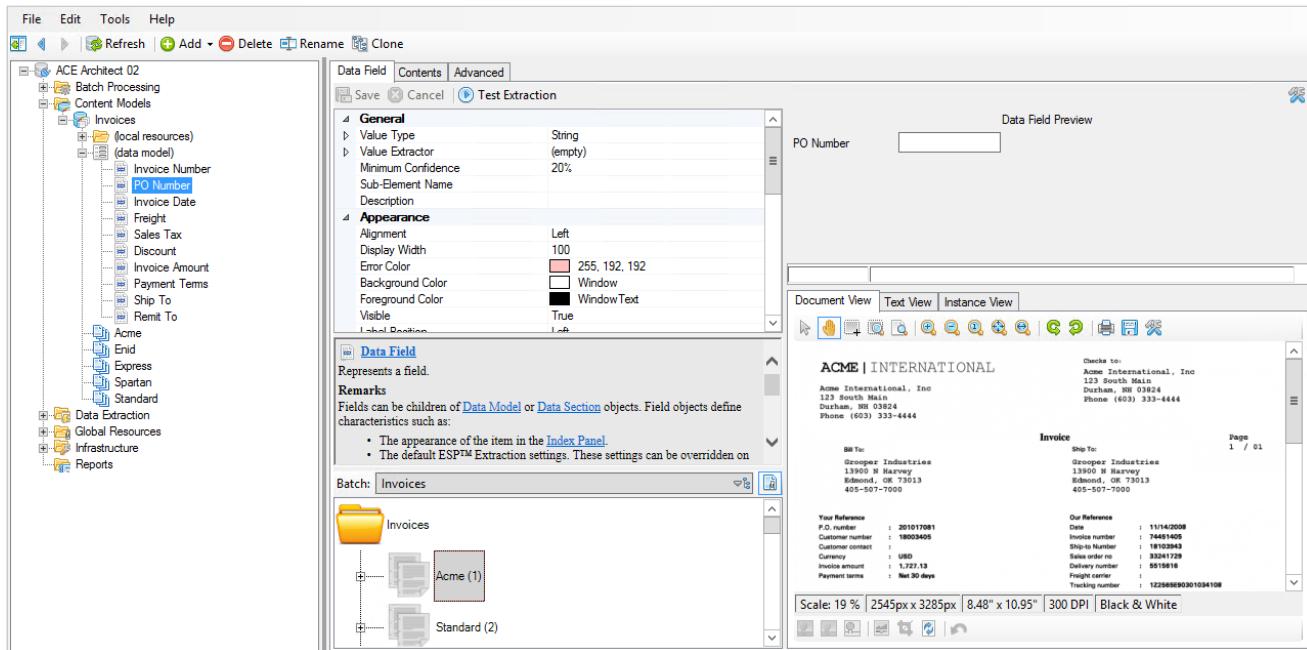
Feature Occurrences: 3 | Total Classes: 9 | Class Trained Items: 0 | Total Weight: 1.000000

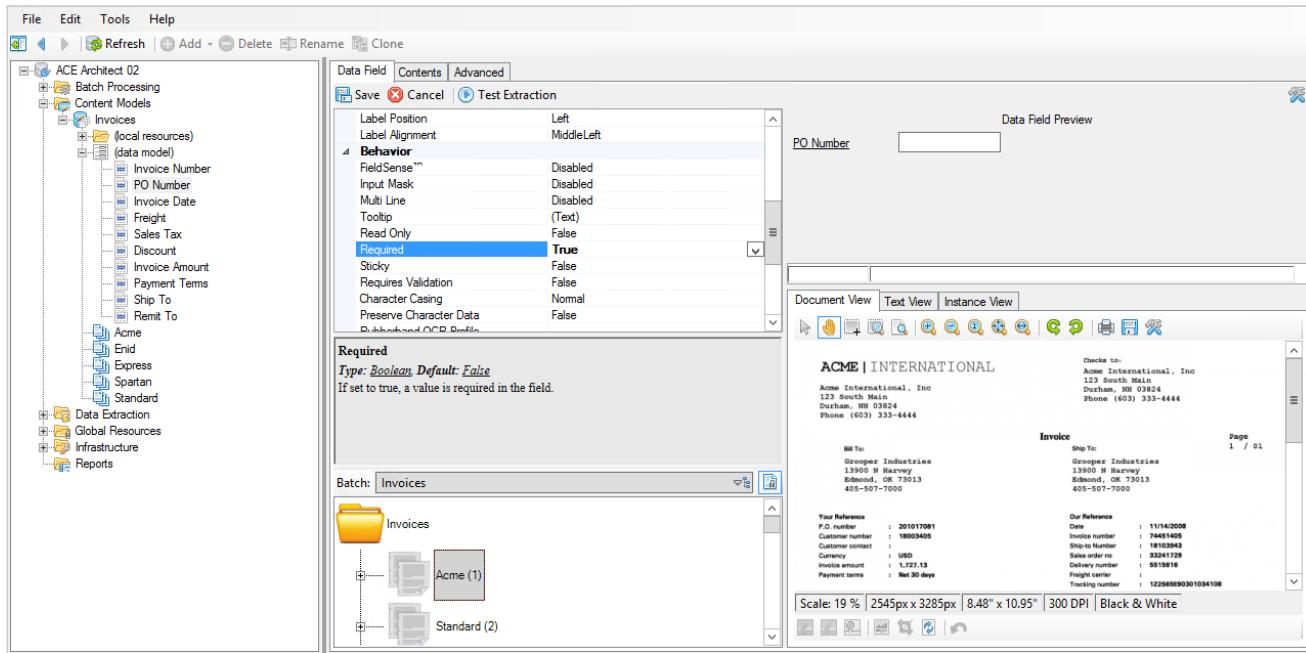
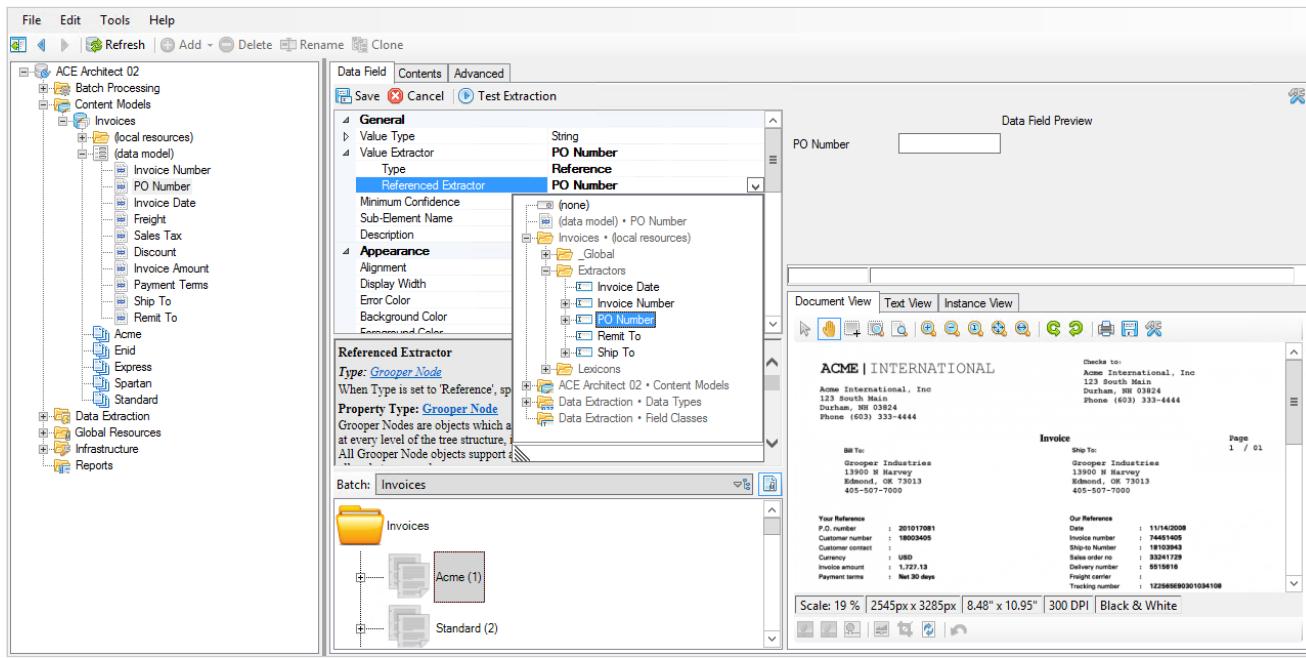
| Feature | Count | CWF | TF | IDF | Weight |
|----------------|-------|-----|----------|----------|----------|
| po number | 1 | 1 | 0.333333 | 1.000000 | 0.333333 |
| invoice amount | 1 | 0 | 0.333333 | 1.000000 | 0.333333 |
| due date | 1 | 0 | 0.333333 | 1.000000 | 0.333333 |



Step

1. In the Data Model, select the PO Number Data Field
2. In the General section, set the Value Extractor > Type property to Reference
3. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > PO Number Field Class.
4. In the Behavior section, set the Required property to True

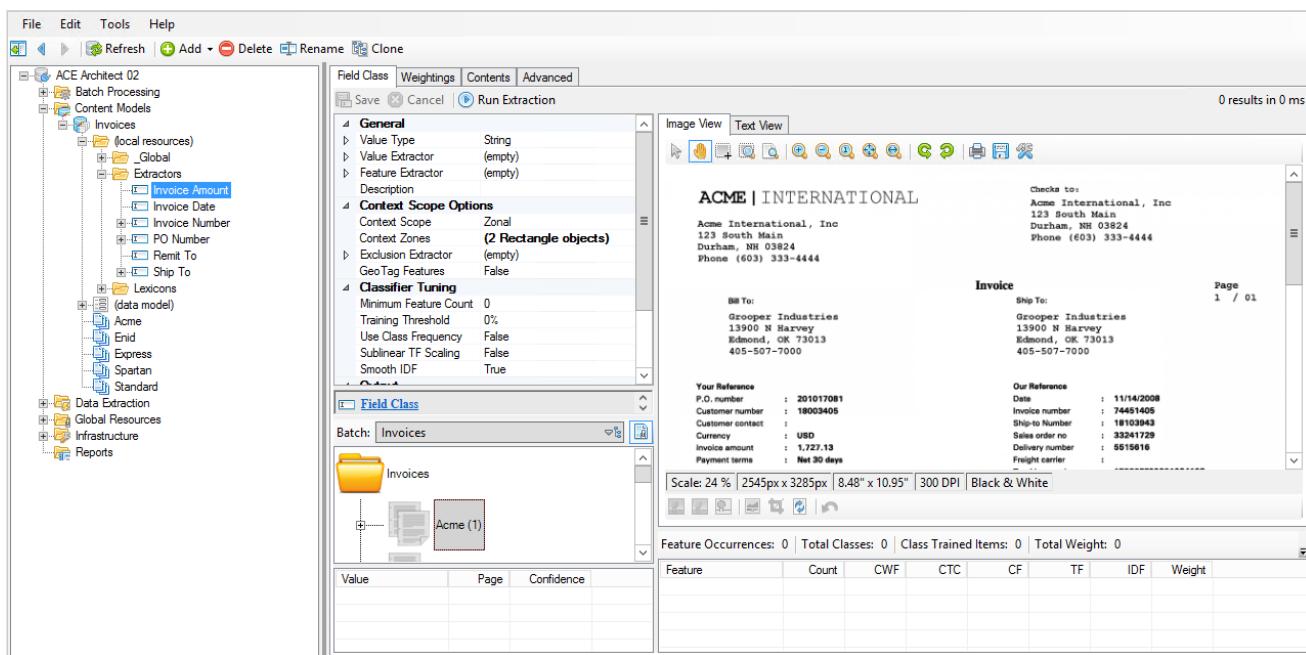
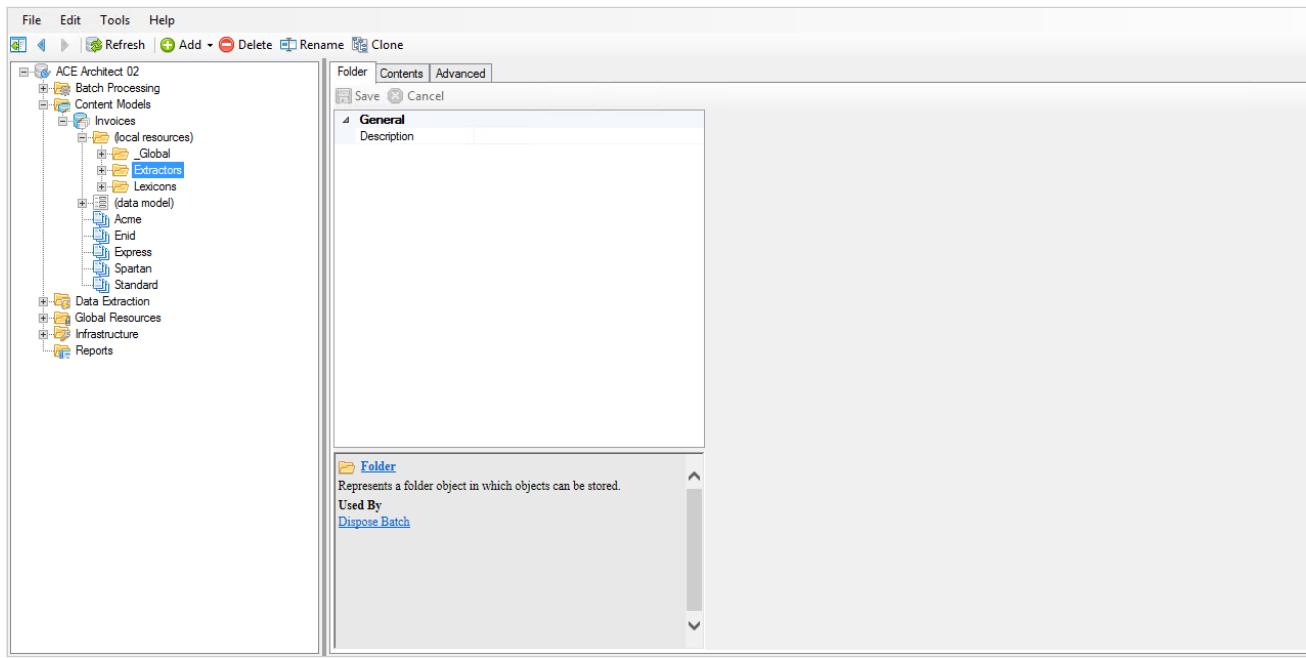




Invoice Amount

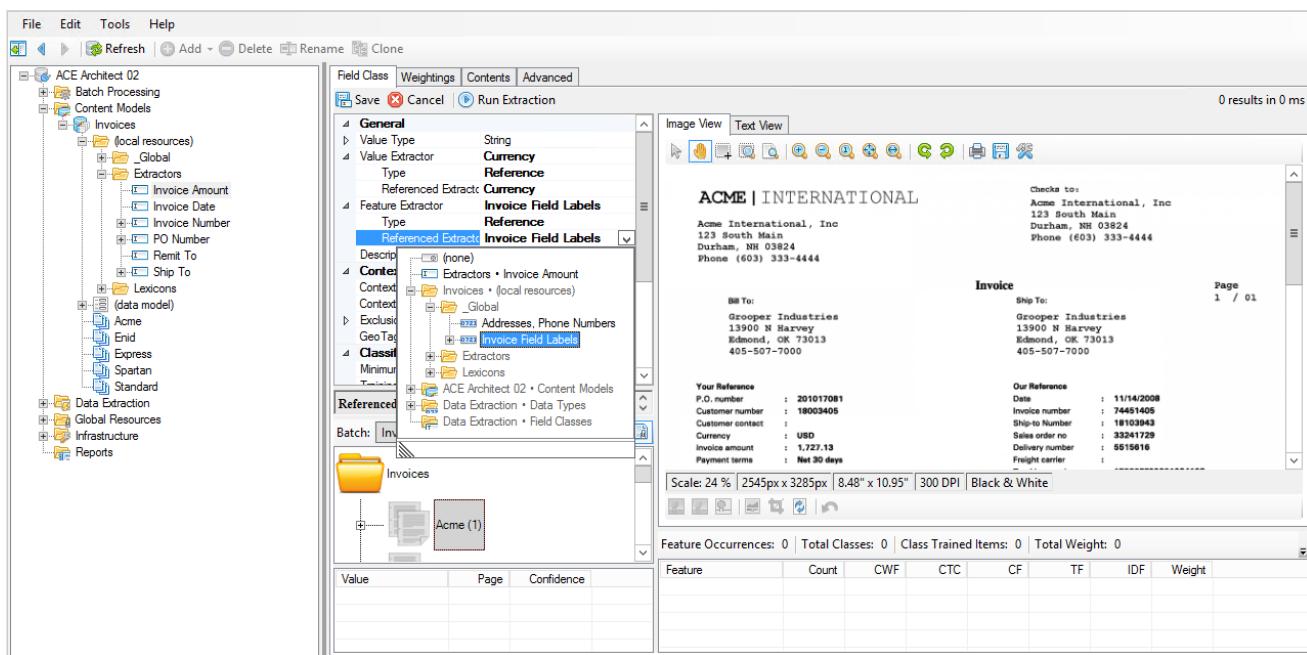
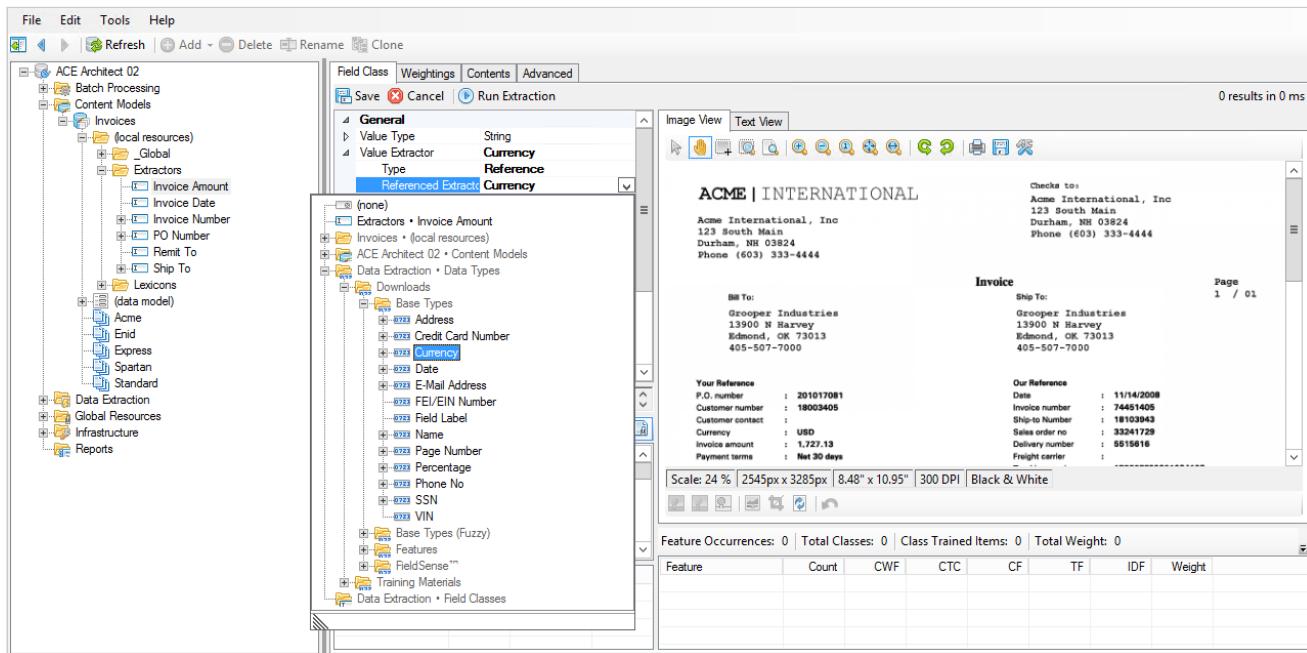
Step

1. Right click on the Extractors folder and select Add > Field Class...
2. Name it **Invoice Amount**



Step

1. In the General section, set the Value Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Data Extraction • Data Types > Downloads > Base Types > **Currency Data Type**
3. Set the Feature Extractor > Type property to Reference
4. Set the Referenced Extractor property to the Invoices • (local resources) > _Global > Invoice Field Labels Data Type



Step

1. Select Acme (1)
2. Click Run Extraction
3. In the candidate list, find the value that has the appropriate feature(s) to train
4. Right click and select Train As Positive

The screenshot shows the ACE Architect 02 interface with the following details:

- File, Edit, Tools, Help** menu bar.
- Toolbar:** Refresh, Add, Delete, Rename, Clone.
- Left Sidebar:** Content Models, Invoices (selected), Lexicons (data model), Acme, Express, Spartan, Standard, Data Extraction, Global Resources, Infrastructure, Reports.
- Central Area:**
 - Field Class:** Weightings, Contents, Advanced.
 - Save, Cancel, Run Extraction** buttons.
 - Image View:** Text View, Image View (selected).
 - General:** Value Type String, Value Extractor Currency Reference (Referenced Extract: Currency).
 - Feature Extractor:** Invoice Field Labels Type Reference (Referenced Extract: Invoice Field Labels).
 - Context Scope Options:** Context Scope Zonal, Context Zones (2 Rectangle objects).
 - Exclusion Extractor:** (empty).
 - GeoTag Features:** False.
 - Classifier Tuning:** Minimum Feature Count 0.
- Preview Area:** Shows an image of an invoice with extracted fields:
 - Bill 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: 74401405, Shipment number: 18103843, Sales order no: 33241728, Delivery number: 5519616, Freight carrier: , Tracking number: 122565E90301034108.
- Bottom Area:** Scale: 24 % | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White.
- Feature Occurrences:** 1 | Total Classes: 0 | Class Trained Items: 0 | Total Weight: 1.000000.
- Feature Table:**

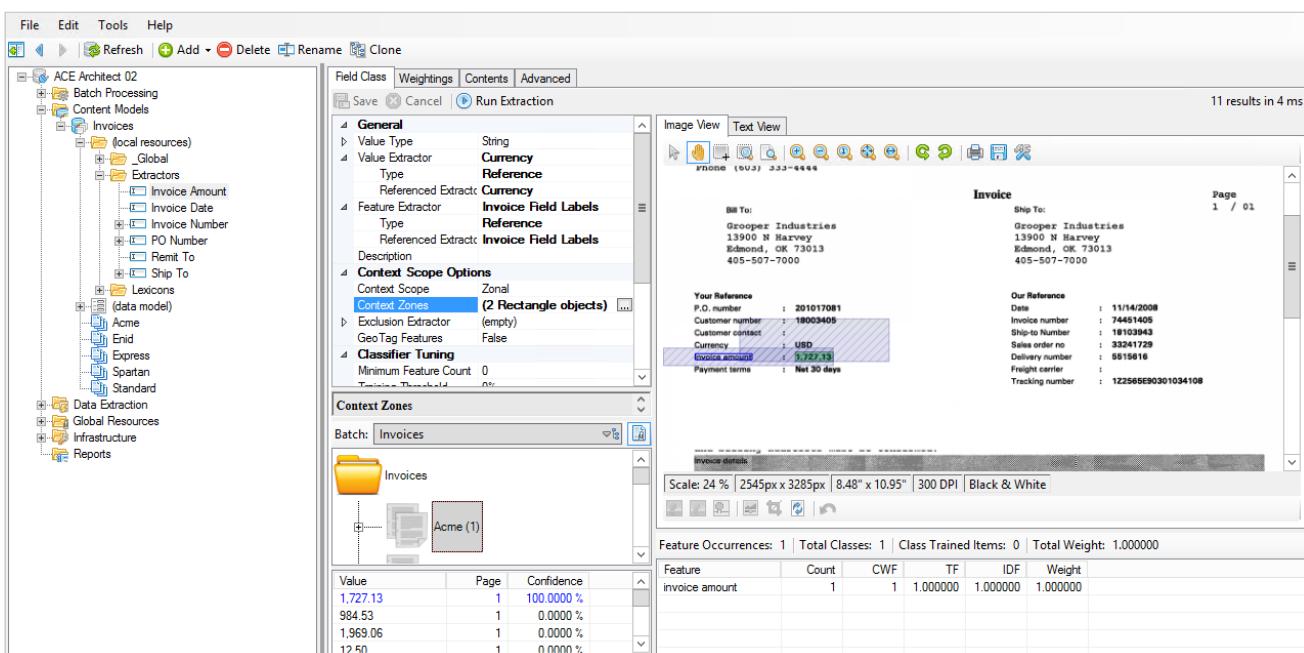
| Feature | Count | CWF | TF | IDF | Weight |
|----------------|-------|-----|----------|----------|----------|
| invoice amount | 1 | 0 | 1.000000 | 1.000000 | 1.000000 |

The screenshot displays the ACE Architect 02 application interface. The left sidebar shows project navigation with sections like Batch Processing, Content Models, and Invoices. The main workspace is divided into several panes:

- Extraction Configuration:** A large pane on the left containing tabs for Field Class, Weightings, Contents, and Advanced. The General tab is active, showing settings for Value Type (String), Value Extractor (Currency), Feature Extractor (Invoice Field Labels), and Context Scope Options (Zonal, Context Zones: (2 Rectangle objects)).
- Preview:** A central pane titled "Image View" showing a scanned document of an invoice. The invoice details include:
 - 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
 - Our Reference: Date : 11/14/2008
 - P.O. number : 18003405
 - Customer number : 18003405
 - Invoice number : 74451405
 - Customer contact : 18003405
 - Sales order no : 33241729
 - Currency : USD
 - Delivery number : 5515616
 - GeoTag Features: False
 - Invoice amount : 1,727.13
 - Payment terms : Net 30 days
- Results:** A bottom-right pane showing feature occurrences and weights. It lists "Feature Occurrences: 1 | Total Classes: 1 | Class Trained Items: 0 | Total Weight: 1.000000". A table shows the extracted value "1,727.13" with a confidence of 100.0000 %.

Step

1. Click to edit the Context Zones property
2. In the Context Zones window, edit the dimensions as follows:
 - a. Zone 1:
 - i. Left: **-4**
 - ii. Top: **-0.05**
 - iii. Right: **0.2**
 - iv. Bottom: **0.1**
 - b. Zone 2:
 - i. Left: **-0.58**
 - ii. Top: **-0.36**
 - iii. Right: **0.3**
 - iv. Bottom: **0.05**
3. Click OK
4. In the Output section, set the Minimum Confidence property to **70%**



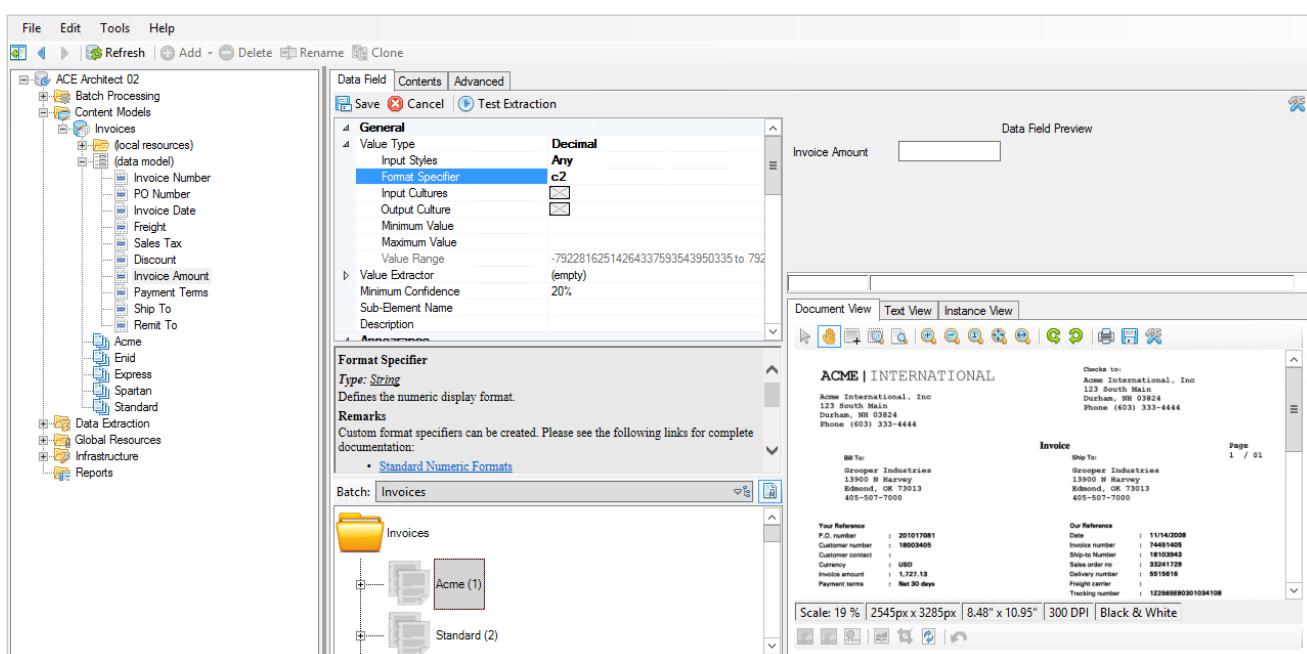
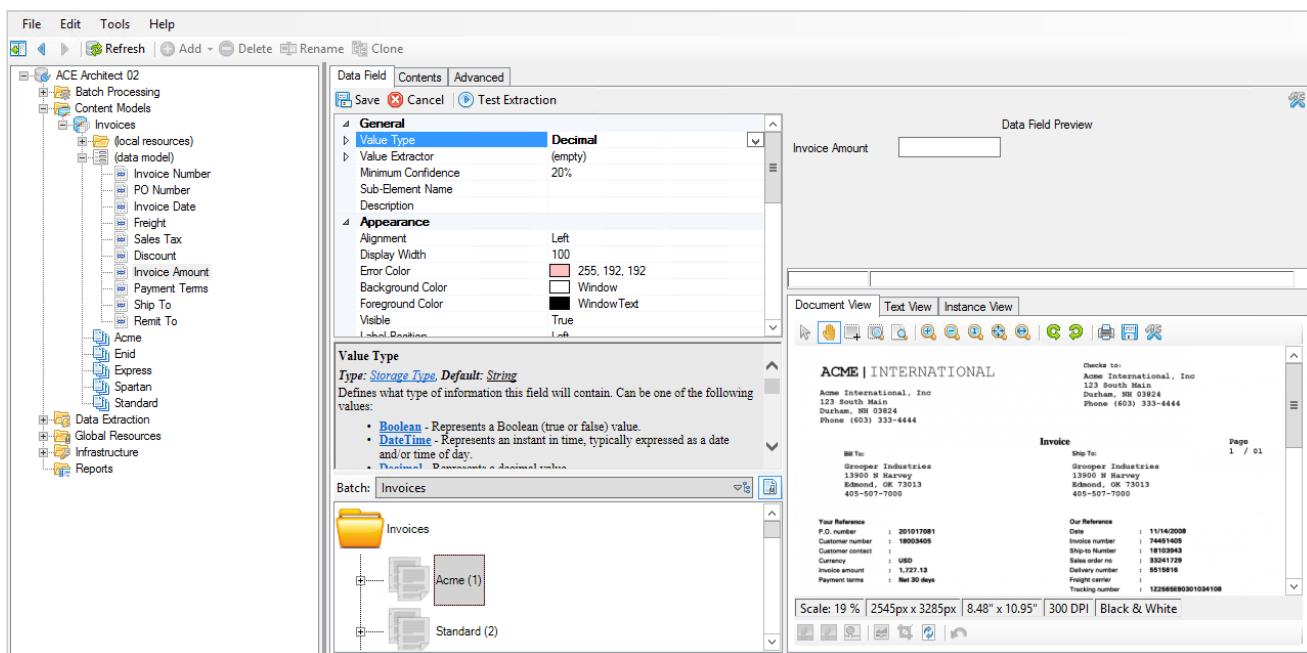
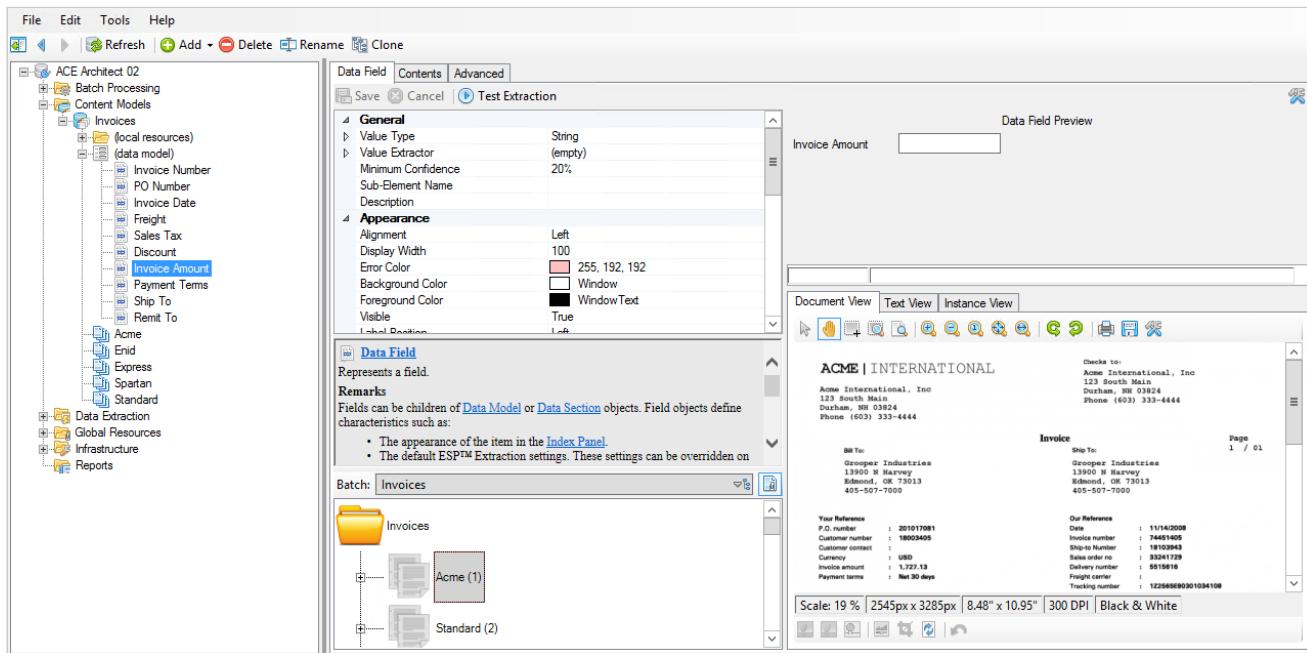
The screenshot shows the ACE Architect 02 interface. On the left, the navigation pane includes sections like Batch Processing, Content Models, Invoices, Extractors, Lexicons, and Reports. The main area displays the 'Field Class' configuration for an 'Invoices' field. Under the 'General' section, the 'Value Type' is set to 'String' and 'Extractor' is set to 'Currency'. A 'Text View' tab is active, showing a sample document snippet and context zones for 'Zone 1' and 'Zone 2'. The 'Zone 1' configuration is '(4,-0.05)(0.2,0.1)'. The 'Zone 2' configuration is '(-1,-0.5):(1,0)'. To the right, a detailed view of the extracted data is shown, including fields like 'Date', 'Invoice number', 'Sales order no', 'Delivery number', 'Freight carrier', and 'Tracking number'. Below this, a feature occurrence table lists the extracted value '1,727.13' with a confidence of 100.0000%.

This screenshot shows the same interface as above, but with different zone configurations. The 'Zone 1' configuration is now '(4,-0.05)(0.2,0.1)' and the 'Zone 2' configuration is '(-0.58,-0.36)(0.3,0.05)'. The rest of the interface and extracted data remain similar to the first screenshot.

This screenshot shows the 'Field Class' configuration with a focus on 'Context Scope Options'. The 'Context Scope' is set to 'Zonal' and 'Context Zones' are specified as '(2 Rectangle objects)'. The 'Output' section is configured with a 'Minimum Confidence' of 50%, 'Collation Method' as 'Individual', and 'Order By' as 'Confidence Descending'. The 'Text View' tab shows the extracted data, including the invoice amount '1,727.13' with a confidence of 100.0000%. The right side of the interface displays the extracted invoice details, including the bill-to and ship-to addresses, payment terms, and other financial information.

Step

1. In the Data Model, select the Invoice Amount Data Field
2. In the General section, set the Value Type to **Decimal**
3. Expand Value Type and set the Format Specifier property to **c2**



Step

1. In the General section, set the Value Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > Invoice Amount Field Class.
3. In the Behavior section, set the Required property to **True**

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources) (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
- Acme
- End
- Express
- Spartan
- Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **Cancel** **Test Extraction**

General

- Value Type: **Decimal**
- Value Extractor: **(empty)**
- Minimum Confidence: 20%
- Sub-Element Name:
- Description:

Appearance

- Alignment: **Left**
- Display Width: **100**
- Error Color: **255, 192, 192**
- Background Color: **Window**
- Foreground Color: **WindowText**
- Visible: **True**
- Label Position: **Left**

Value Extractor

Type: **Embedded Extractor**

The default extractor to use for this field.

Remarks

The default extractor can be overridden on individual [Content Types](#) using a [Field Profile](#).

Property Type: Embedded Extractor

Batch: **Invoices**

Document View **Text View** **Instance View**

ACME | INTERNATIONAL

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

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

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

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

Page 1 / 01

Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources) (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
- Acme
- End
- Express
- Spartan
- Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **Cancel** **Test Extraction**

General

- Value Type: **Decimal**
- Value Extractor: **Invoice Amount**
- Type: **Referenced Extractor**
- Minimum Confidence:
- Sub-Element Name:
- Description:

Appearance

- Alignment:
- Display Width:
- Error Color:
- Background Color:
- Foreground Color:
- Visible:
- Label Position:

Referenced Extractor

Type: **Grooper Node**

When Type is set to 'Reference', specify the node to reference.

Property Type: Grooper Node

Grooper Nodes are objects which appear at every level of the tree structure, in all Grooper Node objects support a

Batch: **Invoices**

Document View **Text View** **Instance View**

ACME | INTERNATIONAL

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

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

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

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

Page 1 / 01

Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources) (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
- Acme
- End
- Express
- Spartan
- Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **Cancel** **Test Extraction**

Label Position: Left

Label Alignment: MiddleLeft

Behavior

- FieldSense™: **Disabled**
- Input Mask: **Disabled**
- Tooltip: **(Text)**
- Read Only: **False**
- Required: **True**
- Sticky: **False**
- Requires Validation: **False**
- Preserve Character Data: **False**
- Rubberband OCR Profile:

Database Lookup

Batch: **Invoices**

Document View **Text View** **Instance View**

ACME | INTERNATIONAL

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

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

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

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

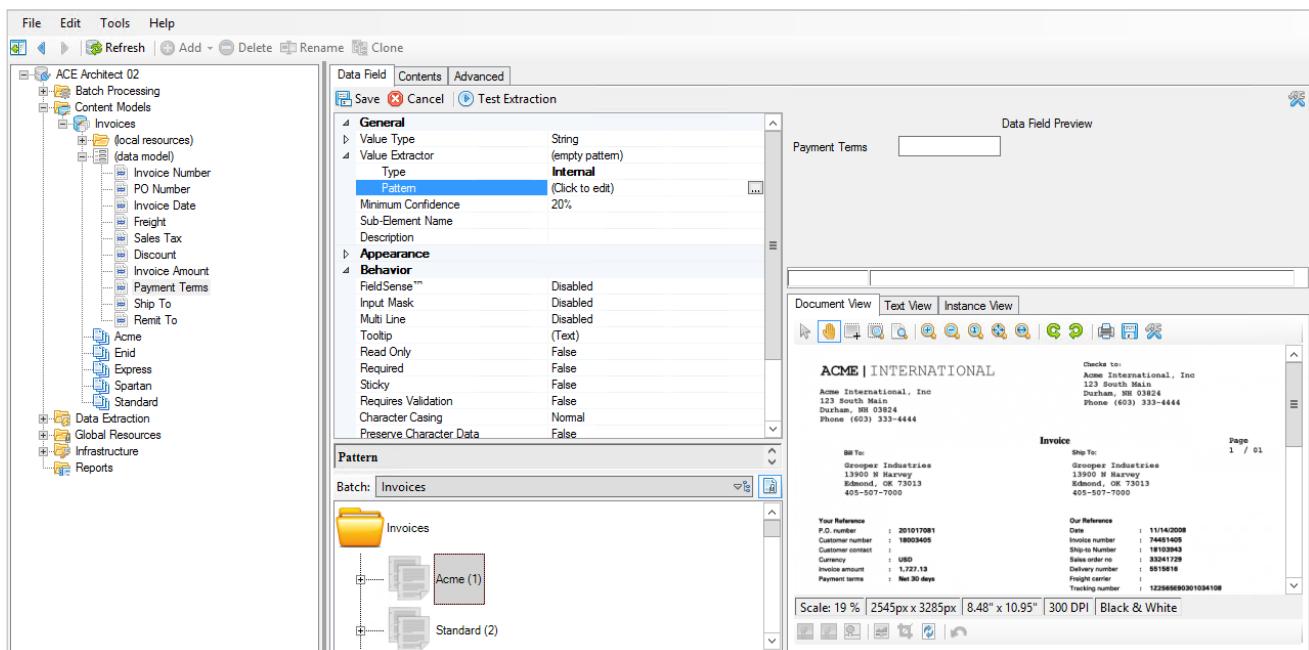
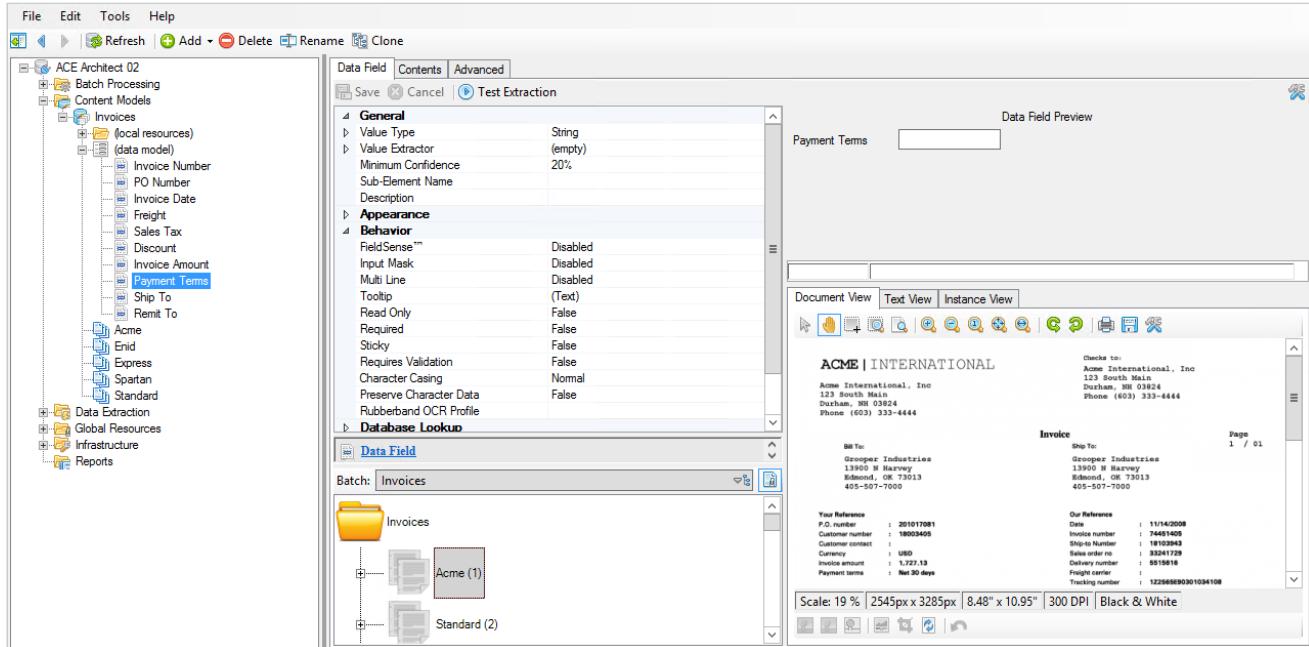
Page 1 / 01

Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

Payment Terms

Step

1. In the Data Model, select the PO Number Data Field
2. In the General section, set the Value Extractor > Type property to Internal
3. Click to edit the pattern



Step

1. Select Acme (16)

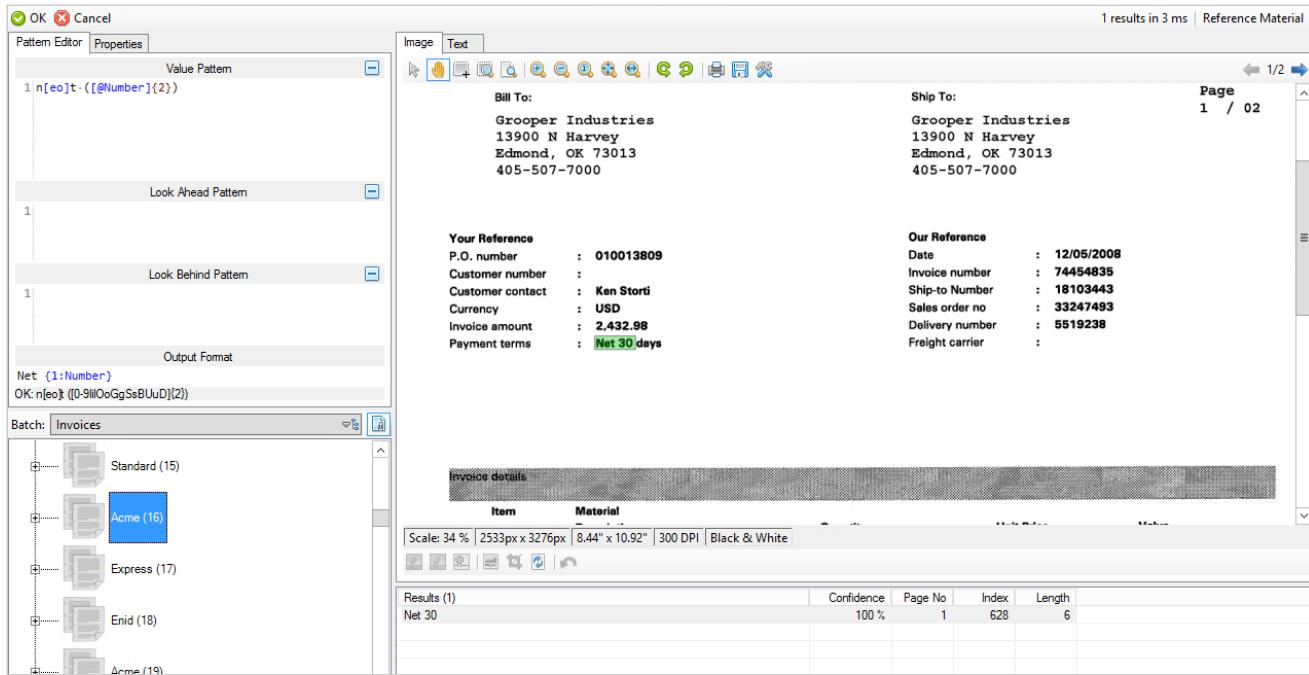
2. In the Value Pattern editor, type:

`n[eo]t ([@Number]{2})`

3. In the Output Format editor, type:

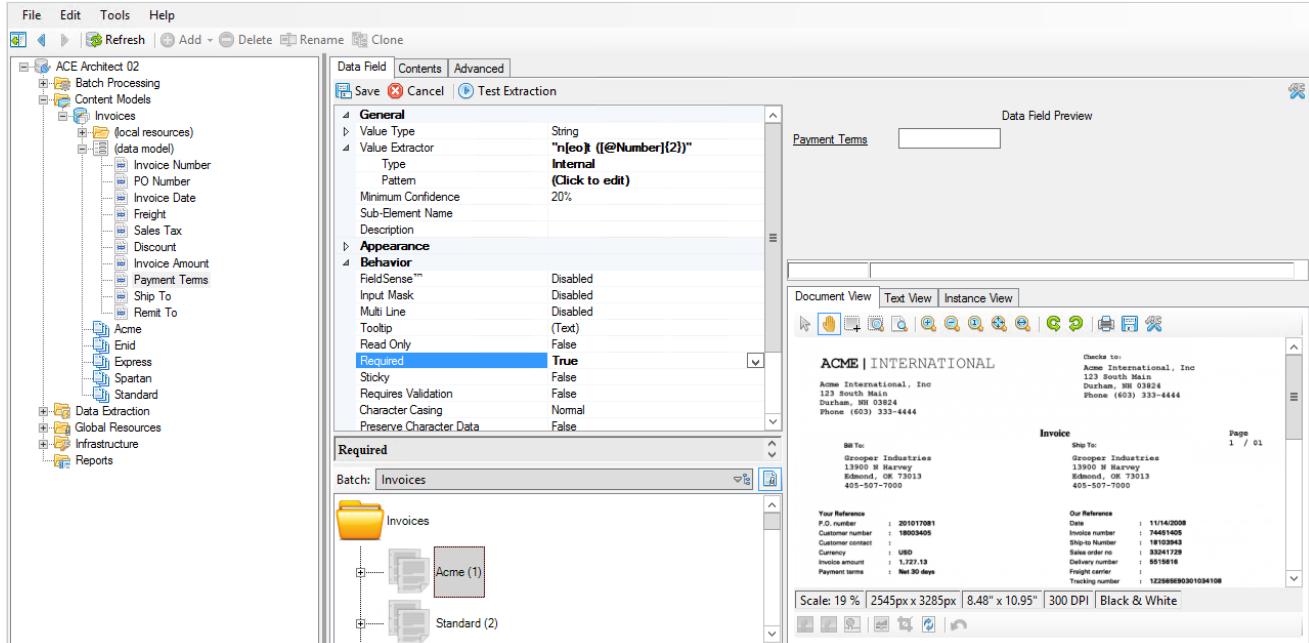
`Net {1:Number}`

4. Click OK



Step

1. On the Payment Terms Data Field, set the Required Property to True

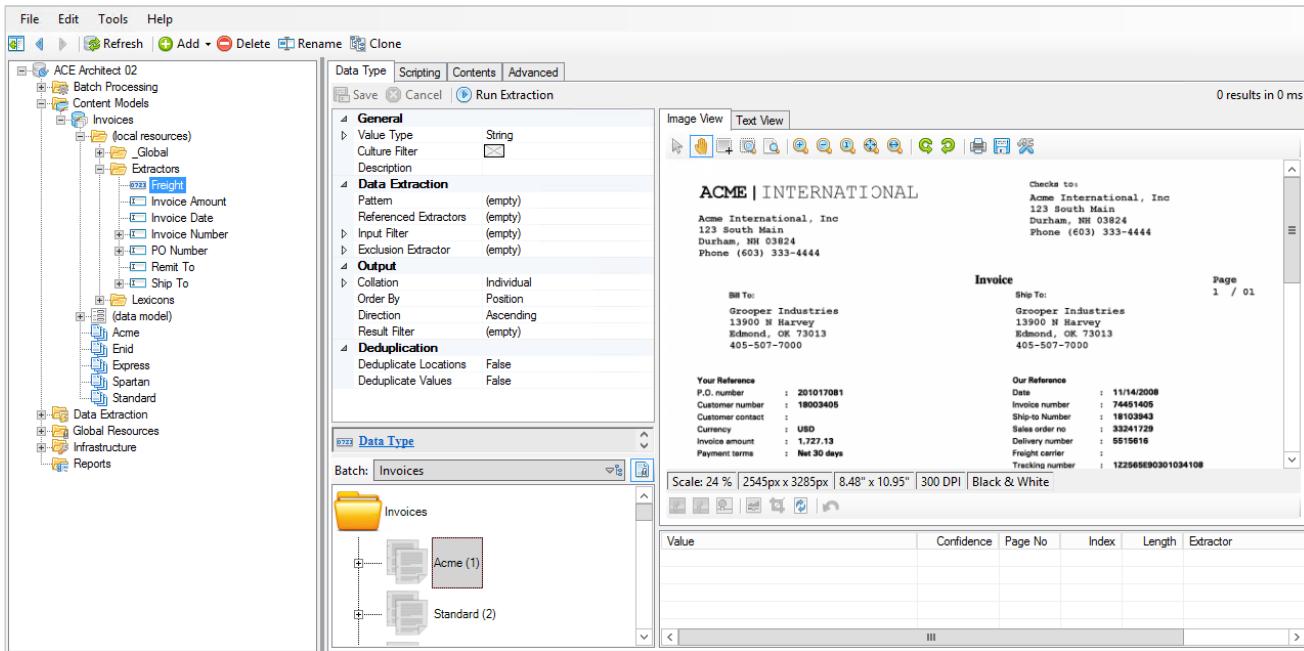
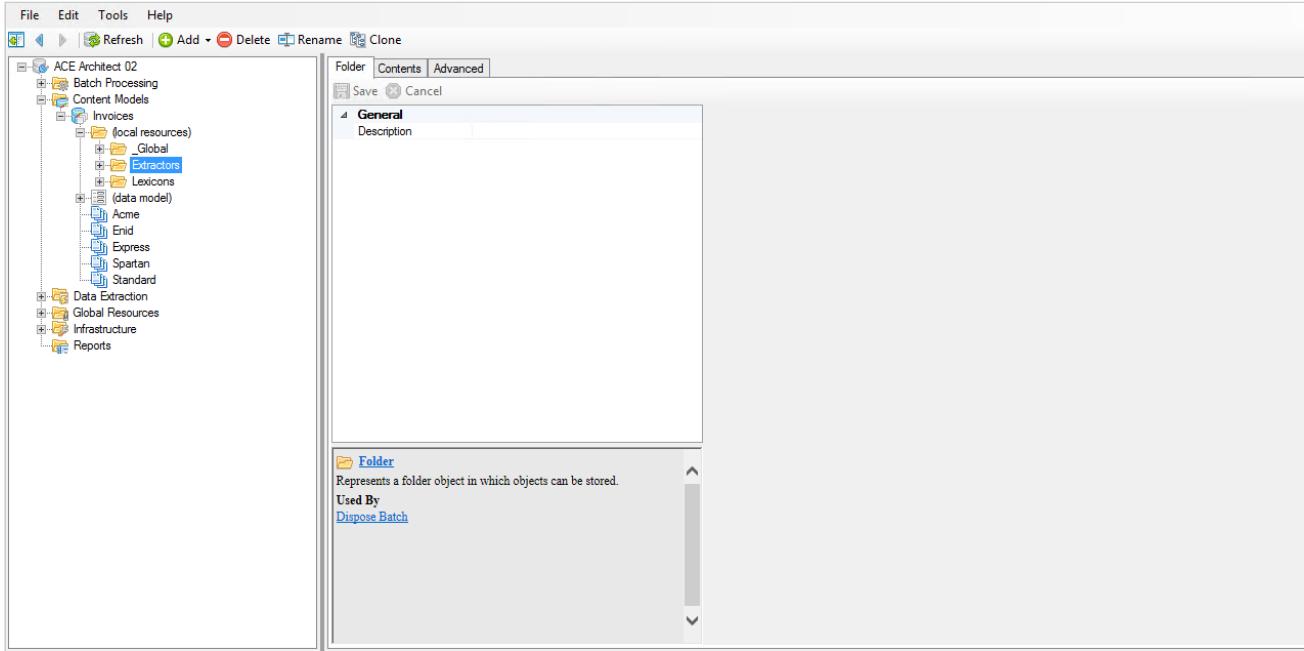


Freight

Step

1. Right click on the Extractors folder and select Add > Data Type...

2. Name it Freight



Step

1. As child extractors under the Freight Data Type, create:

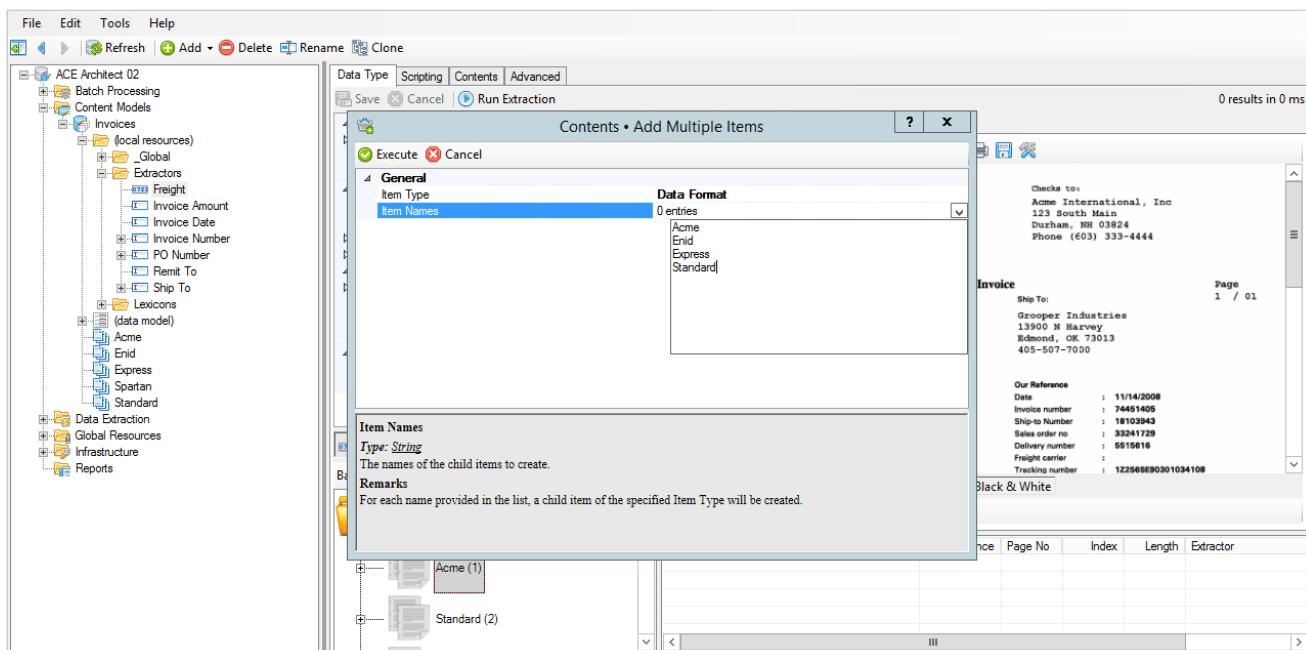
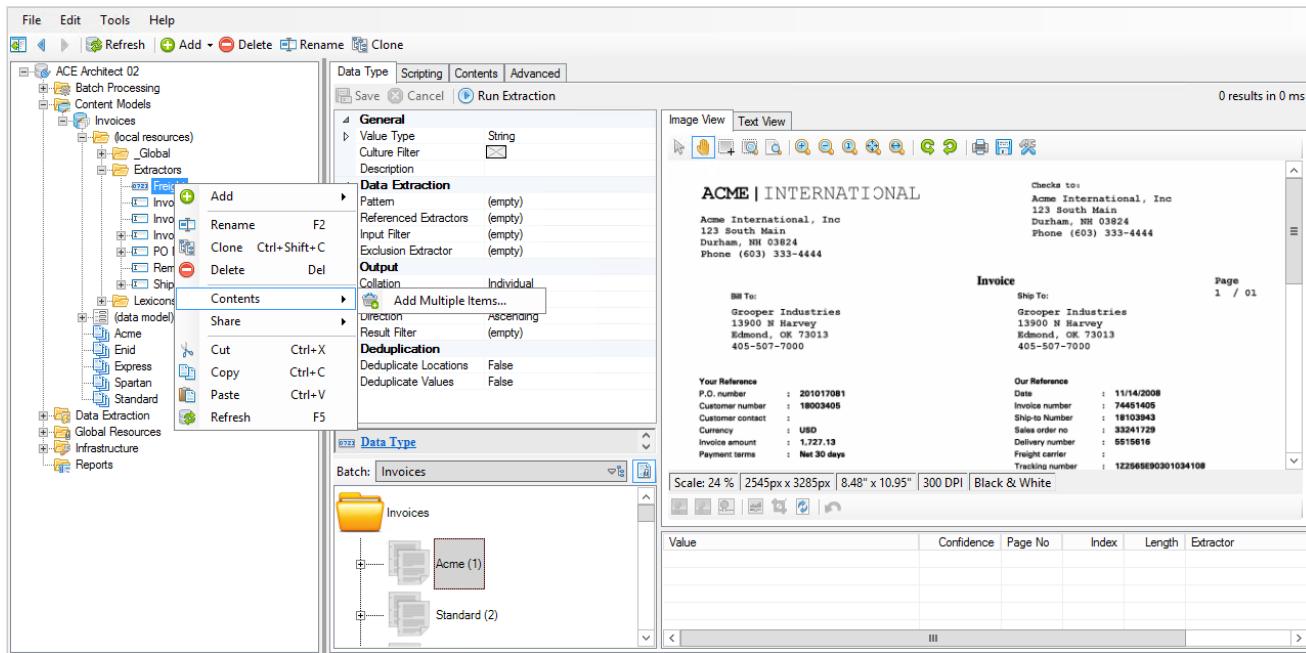
a. Data Format: Acme

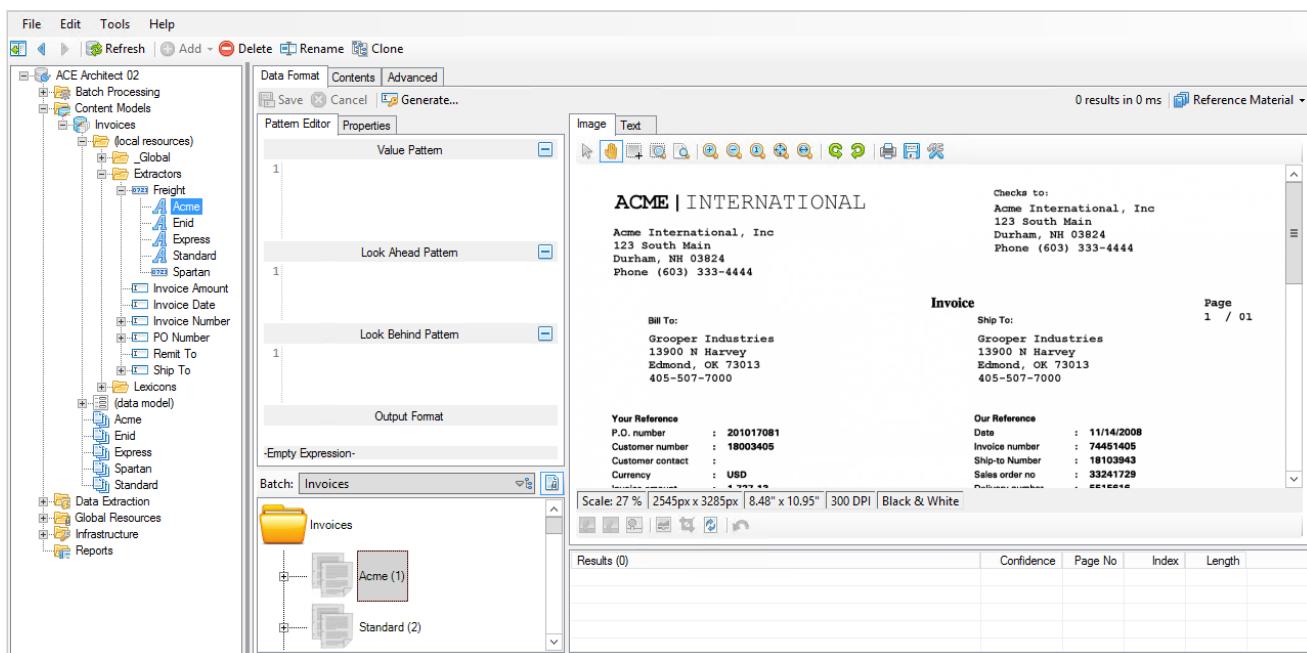
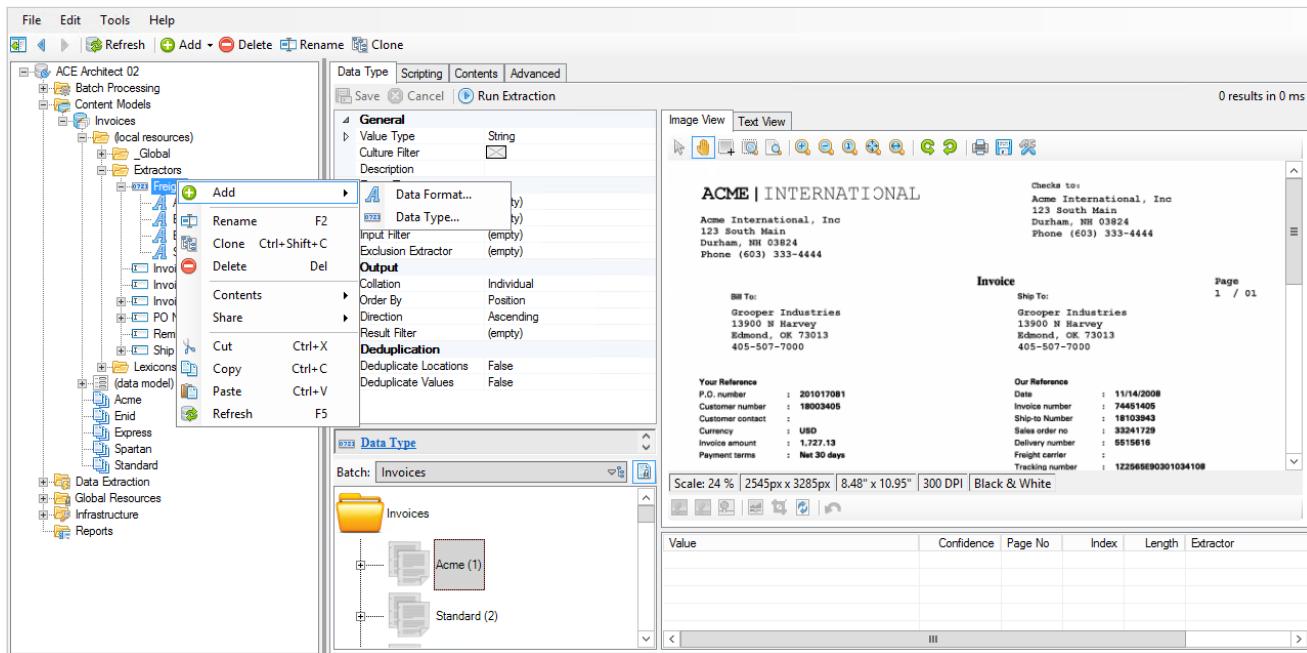
b. Data Format: Enid

c. Data Format: Express

d. Data Format: Standard

e. Data Type: Spartan





--- Edit the patterns for each of the Data Formats: - Acme

Make sure to enable Tab Marking

```
```tab="Value Pattern"
(@Number.{3,12})
```

```

```
```tab="Look Ahead Pattern"
WS.FREIGHT[0o]231[^n]+n
```

```

```
[^n]+\n
[^t]+\t
[^t]+\v
```

```

```
```tab="Look Behind Pattern"
\r\n
```

```

```
```tab="Output Format"
{1:Number}
```

```

- Enid

```
```tab="Value Pattern"
(@Number.{3,12})
```

```

```tab="Look Ahead Pattern"

020-0027[^\\n]*?\\n[^\\r]*?

3

```tab="Look Behind Pattern"

\r\n

3

```tab="Output Format"

{1:Number}

1

- Express

```tab="Value Pattern"

([@Number.] {3,12})

10

```tab="Look Ahead Pattern"

shipping charge:?[^\0-9]{0,4}

1

```tab="Output Format"

{1:Number}

3

#### - Standard

```tab="Value Pattern"

([@Number.] {3,12})

1

```tab="Look Ahead Pattern"

shipping charges

3

```tab="Output Format"

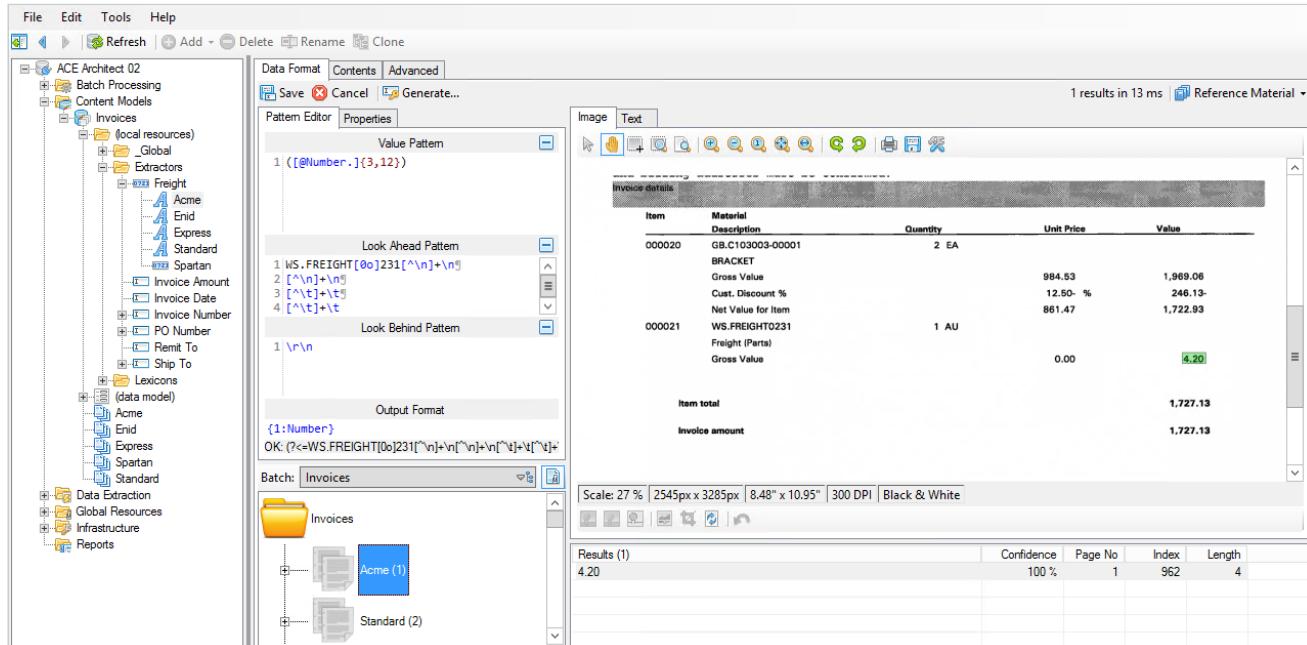
{1:Number}

1

```tab="0u

{1 :Number}

3



**ACE Architect 02**

- Batch Processing
- Content Models
- Invoices
  - (local resources)
  - Global
  - Extractors
    - Freight
      - Acme
      - Enid
      - Express
      - Standard
      - Spartan
      - Invoice Amount
      - Invoice Date
      - Invoice Number
      - PO Number
      - Remit To
      - Ship To
  - Lexicons (data model)
    - Acme
    - Enid
    - Express
    - Spartan
    - Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

**Data Format** [Save] [Cancel] [Generate...]

**Pattern Editor** [Properties]

**Value Pattern**

```
1 (@Number.){3,12})
```

**Look Ahead Pattern**

```
1 020-0027["\n"]*?\n[^r]*?
```

**Look Behind Pattern**

```
1 \r\n
```

**Output Format**

```
{1:Number}
OK: (?<=020-0027["\n"]*?\n[^r]*?)([0-9]{3}OoGgSsBuuD){3,1}
```

**Batch:** Invoices

- Acme (9)
- Enid (10)
- Acme (11)

**Image** [Text]

TO: Edmond, OK 73013 TO: Edmond, OK 73013

\*\*\*\*\*

| ORDER NUMBER            | ORDER DATE       | CUSTOMER NUMBER  | SALES PERSON | PURCHASE ORDER NUMBER                               | SHIP VIA | SHIP DATE      | TERMS  |
|-------------------------|------------------|------------------|--------------|-----------------------------------------------------|----------|----------------|--------|
| 100050168               | 12-02-08         | CUS001605        | 099321       | 201017532                                           | FX1      | 12-02-08       | Net 30 |
| <b>ITEM DESCRIPTION</b> |                  |                  |              |                                                     |          |                |        |
| ITEM NUMBER             | QUANTITY ORDERED | QUANTITY SHIPPED | STK UNIT     | PRICE UNIT                                          | DISCOUNT | EXTENDED PRICE |        |
| 7710.1262.3             | 1.0000           | 1.0000           | D            | 100093070<br>Block, Mtg., Hot Head<br>238.680 Gross |          | 238.680        |        |
| 020-0027                | 1.0000           | 1.0000           | ea           | Shipping and Handling<br>60.0000 Gross              | 60.000   |                |        |
|                         |                  |                  |              | Order Discount Amount                               | -52.510  |                |        |

Scale: 27 % | 2548px x 3288px | 8.49" x 10.96" | 300 DPI | Black & White

Results (1) | Confidence | Page No | Index | Length

60.000 | 100 % | 1 | 743 | 6

**ACE Architect 02**

- Batch Processing
- Content Models
- Invoices
  - (local resources)
  - Global
  - Extractors
    - Freight
      - Acme
      - Enid
      - Express
      - Standard
      - Spartan
      - Invoice Amount
      - Invoice Date
      - Invoice Number
      - PO Number
      - Remit To
      - Ship To
  - Lexicons (data model)
    - Acme
    - Enid
    - Express
    - Spartan
    - Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

**Data Format** [Save] [Cancel] [Generate...]

**Pattern Editor** [Properties]

**Value Pattern**

```
1 (@Number.){3,12})
```

**Look Ahead Pattern**

```
1 shipping-charge?:[^d]{0,4}
```

**Look Behind Pattern**

```
1 \r\n
```

**Output Format**

```
{1:Number}
OK: (?<shipping-charge?:[^d]{0,4})([0-9]{3}OoGgSsBuuD){3,1}
```

**Batch:** Invoices

- Enid (5)
- Express (6)

**Image** [Text]

Kalamazoo, MI 49003  
405-507-7000

Vendor Number: M00000032  
Customer Account Number: 138463500

| Line | McMaster Part Number | Description/Customer ID #                    | You Ordered | We Shipped | Balance Due | Unit Price | Extended Amount |
|------|----------------------|----------------------------------------------|-------------|------------|-------------|------------|-----------------|
| 1    | 4192A64              | CARBIDE-TIPPED HOLE SAW, 6" DIAMETER, 1-1/2" | 1           | 1          | 0           | 78.34 EA   | 78.34           |

Merchandise Amount: 78.34  
Shipping Charge: 4.37  
Invoice Total: \$82.71

The material billed on this invoice was shipped as follows:  
McMaster Packing List Number: 1133638-01 Shipped Via: FEDEX GRND PPD on December 2, 2008  
Total Packages: 1 Total Weight: 3 lbs

Scale: 28 % | 2498px x 3287px | 8.33" x 10.96" | 300 DPI | Black & White

Results (1) | Confidence | Page No | Index | Length

4.37 | 100 % | 1 | 807 | 4

**ACE Architect 02**

- Batch Processing
- Content Models
- Invoices
  - (local resources)
  - Global
  - Extractors
    - Freight
      - Acme
      - Enid
      - Express
      - Standard
      - Spartan
      - Invoice Amount
      - Invoice Date
      - Invoice Number
      - PO Number
      - Remit To
      - Ship To
  - Lexicons (data model)
    - Acme
    - Enid
    - Express
    - Spartan
    - Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

**Data Format** [Save] [Cancel] [Generate...]

**Pattern Editor** [Properties]

**Value Pattern**

```
1 (@Number.){3,12})
```

**Look Ahead Pattern**

```
1 shipping-charge:s
```

**Look Behind Pattern**

```
1
```

**Output Format**

```
{1:Number}
OK: (?<shipping-charge:s)([0-9]{3}OoGgSsBuuD){3,12})
```

**Batch:** Invoices

- Acme (1)
- Standard (2)

**Image** [Text]

NUMBER OF PKGS: 0 WEIGHT: 3.20  
DATE SHIPPED: 1/11/2008  
SHIPPING: UPS GROUND REGIONAL  
TRACKING NO: 12309870328242140

These items are sold for domestic consumption in the United States. If exported, purchaser assumes full responsibility for compliance with US export controls.

PAYMENT TERMS NET 30 DAYS. PAY THIS INVOICE NO STATEMENT SENT PAYABLE IN U.S. DOLLARS.

**AMOUNT DUE 51.17**

**PLEASE DETACH THIS PORTION AND RETURN WITH YOUR PAYMENT.**

Scale: 27 % | 2545px x 3274px | 8.48" x 10.91" | 300 DPI | Black & White

Results (1) | Confidence | Page No | Index | Length

7.71 | 100 % | 1 | 688 | 4

## Step

1. Create two Data Formats as children of the Spartan Data Type:

a. Freight

b. Value 1

2. Edit the Value Patterns for each of the Data Formats:

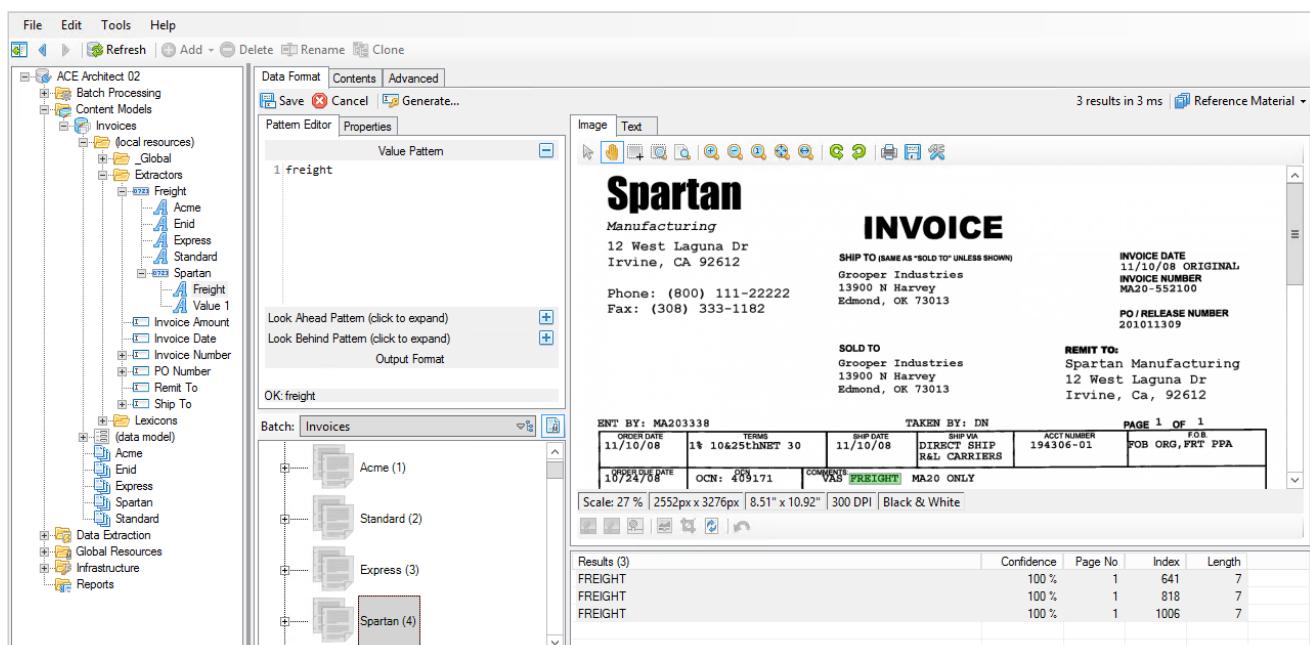
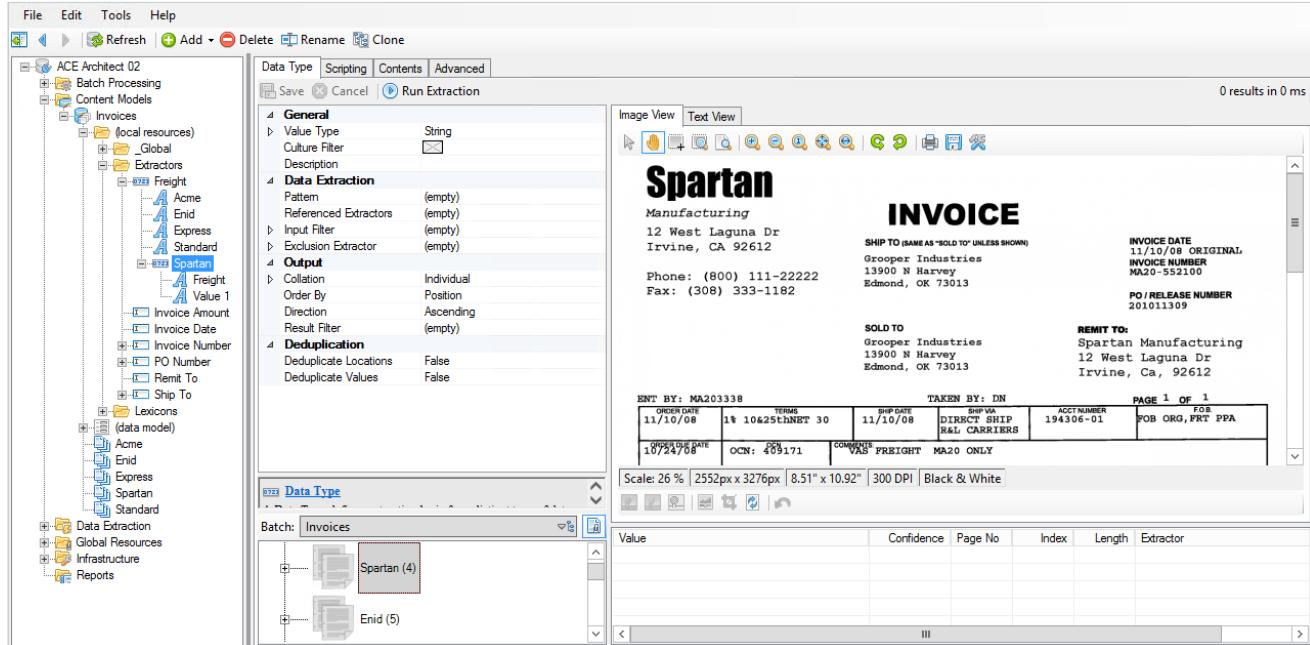
Freight

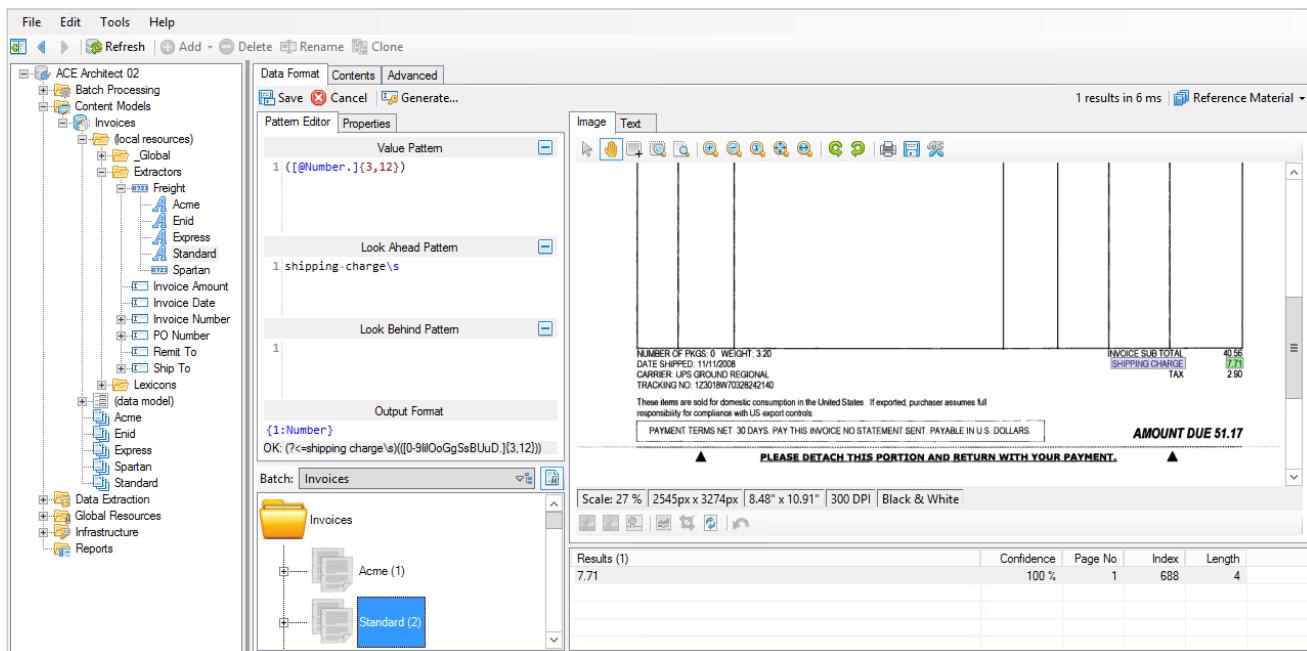
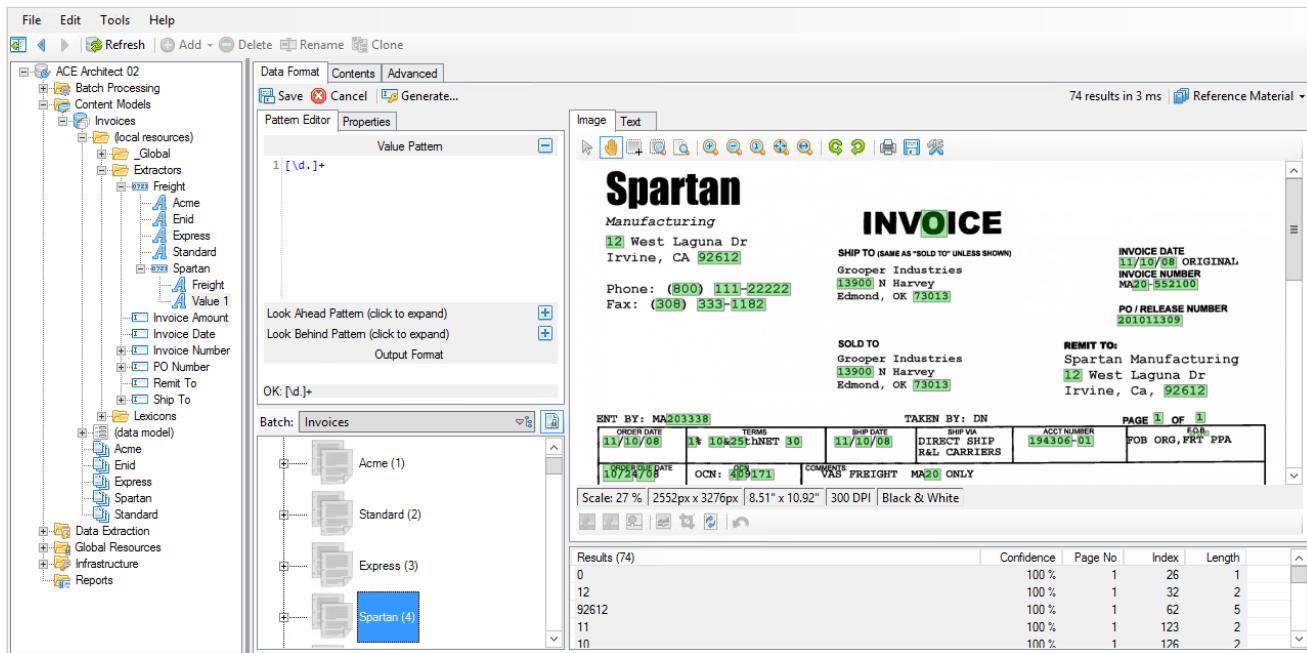
**freight**

Value 1

3. Right click on the Value 1 Data Format and select Clone

4. Name the clone **Value 2**





## Step

1. Select the Spartan Data Type
2. Set the Collation property to **Ordered Array**
3. Expand Collation and set the Vertical Layout property to **Enabled**
4. Expand the Vertical Layout property
5. Set the Maximum Distance property to **0.25**
6. Set the Combine Method property to **Sum**

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
  - Invoices
    - (local resources)
    - \_Global
    - Extractors
      - Freight
      - Acme
      - Enid
      - Express
      - Standard
      - Spartan
        - Freight
        - Value 1
        - Value 2
      - Invoice Amount
      - Invoice Date
      - PO Number
      - Remit To
      - Ship To
  - Lexicons (data model)
    - Acme
    - Enid
    - Express
    - Spartan
    - Standard
  - Data Extraction
  - Global Resources
  - Infrastructure
  - Reports

Data Type Scripting Contents Advanced

Save Cancel Run Extraction

**General**

  - Value Type String
  - Culture Filter
  - Description

**Data Extraction**

  - Pattern (empty)
  - Referenced Extractors (empty)
  - Input Filter (empty)
  - Exclusion Extractor (empty)

**Output**

  - Collation Individual
  - Order By Position
  - Direction Ascending
  - Result Filter (empty)

**Deduplication**

  - Duplicate Locations False
  - Duplicate Values False

Image View Text View

**Spartan**

**INVOICE**

SHIP TO (SAME AS "REMIT TO" UNLESS SHOWN)

Grooper Industries  
13900 N Harvey  
Edmond, OK 73013

Phone: (800) 111-22222  
Fax: (508) 333-1182

INVOICE DATE 11/10/08 ORIGINAL  
INVOICE NUMBER MA20-1552109  
PO / RELEASE NUMBER 201011309

SOLD TO Grooper Industries  
13900 N Harvey  
Edmond, OK 73013

REMIT TO Spartan Manufacturing  
12 West Laguna Dr  
Irvine, Ca, 92612

TAKEN BY: DN PAGE 1 OF 1

|                     |                    |                           |                                    |                       |                  |
|---------------------|--------------------|---------------------------|------------------------------------|-----------------------|------------------|
| ORDER DATE 11/10/08 | TERMS 10-30 NET 30 | SHIP DATE 11/10/08        | SHIP VIA DIRECT SHIP & L. CARRIERS | ACCT NUMBER 194306-01 | POB ORG, FRT PPA |
| 10/27/08            | OCN: 123-456-7890  | COMMOD FREIGHT MA-20 ONLY |                                    |                       |                  |

Scale: 26% | 2552px x 3276px | 8.51" x 10.92" | 300 DPI | Black & White

Results (151) Confidence Page No Index Length Extractor

|    |      |   |    |   |         |
|----|------|---|----|---|---------|
| 0  | 100% | 1 | 26 | 1 | Value 1 |
| 0  | 100% | 1 | 26 | 1 | Value 2 |
| 12 | 100% | 1 | 32 | 2 | Value 2 |
| 12 | 100% | 1 | 32 | 2 | Value 1 |
| 0  | 100% | 1 | 63 | 5 | Value 2 |

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
  - Invoices
    - (local resources)
    - \_Global
    - Extractors
      - Freight
      - Acme
      - Enid
      - Express
      - Standard
      - Spartan
        - Freight
        - Value 1
        - Value 2
      - Invoice Amount
      - Invoice Date
      - PO Number
      - Remit To
      - Ship To
  - Lexicons (data model)
    - Acme
    - Enid
    - Express
    - Spartan
    - Standard
  - Data Extraction
  - Global Resources
  - Infrastructure
  - Reports

Data Type Scripting Contents Advanced

Save Cancel Run Extraction

**General**

  - Value Type String
  - Culture Filter
  - Description

**Data Extraction**

  - Pattern (empty)
  - Referenced Extractors (empty)
  - Input Filter (empty)
  - Exclusion Extractor (empty)

**Output**

  - Collation Ordered Array
  - Order By Position
  - Direction Ascending
  - Result Filter (empty)

**Deduplication**

  - Duplicate Locations False
  - Duplicate Values False

Image View Text View

**Spartan**

**INVOICE**

SHIP TO (SAME AS "REMIT TO" UNLESS SHOWN)

Grooper Industries  
13900 N Harvey  
Edmond, OK 73013

Phone: (800) 111-22222  
Fax: (508) 333-1182

INVOICE DATE 11/10/08 ORIGINAL  
INVOICE NUMBER MA20-1552109  
PO / RELEASE NUMBER 201011309

SOLD TO Grooper Industries  
13900 N Harvey  
Edmond, OK 73013

REMIT TO Spartan Manufacturing  
12 West Laguna Dr  
Irvine, Ca, 92612

TAKEN BY: DN PAGE 1 OF 1

|                     |                    |                           |                                    |                       |                  |
|---------------------|--------------------|---------------------------|------------------------------------|-----------------------|------------------|
| ORDER DATE 11/10/08 | TERMS 10-30 NET 30 | SHIP DATE 11/10/08        | SHIP VIA DIRECT SHIP & L. CARRIERS | ACCT NUMBER 194306-01 | POB ORG, FRT PPA |
| 10/27/08            | OCN: 123-456-7890  | COMMOD FREIGHT MA-20 ONLY |                                    |                       |                  |

Scale: 26% | 2552px x 3276px | 8.51" x 10.92" | 300 DPI | Black & White

Results (151) Confidence Page No Index Length Extractor

|    |      |   |    |   |         |
|----|------|---|----|---|---------|
| 0  | 100% | 1 | 26 | 1 | Value 1 |
| 0  | 100% | 1 | 26 | 1 | Value 2 |
| 12 | 100% | 1 | 32 | 2 | Value 2 |
| 12 | 100% | 1 | 32 | 2 | Value 1 |
| 0  | 100% | 1 | 63 | 5 | Value 2 |

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
  - Invoices
    - (local resources)
    - \_Global
    - Extractors
      - Freight
      - Acme
      - Enid
      - Express
      - Standard
      - Spartan
        - Freight
        - Value 1
        - Value 2
      - Invoice Amount
      - Invoice Date
      - PO Number
      - Remit To
      - Ship To
  - Lexicons (data model)
    - Acme
    - Enid
    - Express
    - Spartan
    - Standard
  - Data Extraction
  - Global Resources
  - Infrastructure
  - Reports

Data Type Scripting Contents Advanced

Save Cancel Run Extraction

**General**

  - Value Type String
  - Culture Filter
  - Description

**Data Extraction**

  - Pattern (empty)
  - Referenced Extractors (empty)
  - Input Filter (empty)
  - Exclusion Extractor (empty)

**Output**

  - Collation
    - Horizontal Layout
    - Vertical Layout Enabled
  - Flow Layout
  - Combine Method Individual
  - Result Separator
  - Order By Position
  - Direction Ascending
  - Result Filter (empty)

**Deduplication**

  - Duplicate Locations False
  - Duplicate Values False

Image View Text View

**Spartan**

**INVOICE**

SHIP TO (SAME AS "REMIT TO" UNLESS SHOWN)

Grooper Industries  
13900 N Harvey  
Edmond, OK 73013

Phone: (800) 111-22222  
Fax: (508) 333-1182

INVOICE DATE 11/10/08 ORIGINAL  
INVOICE NUMBER MA20-1552109  
PO / RELEASE NUMBER 201011309

SOLD TO Grooper Industries  
13900 N Harvey  
Edmond, OK 73013

REMIT TO Spartan Manufacturing  
12 West Laguna Dr  
Irvine, Ca, 92612

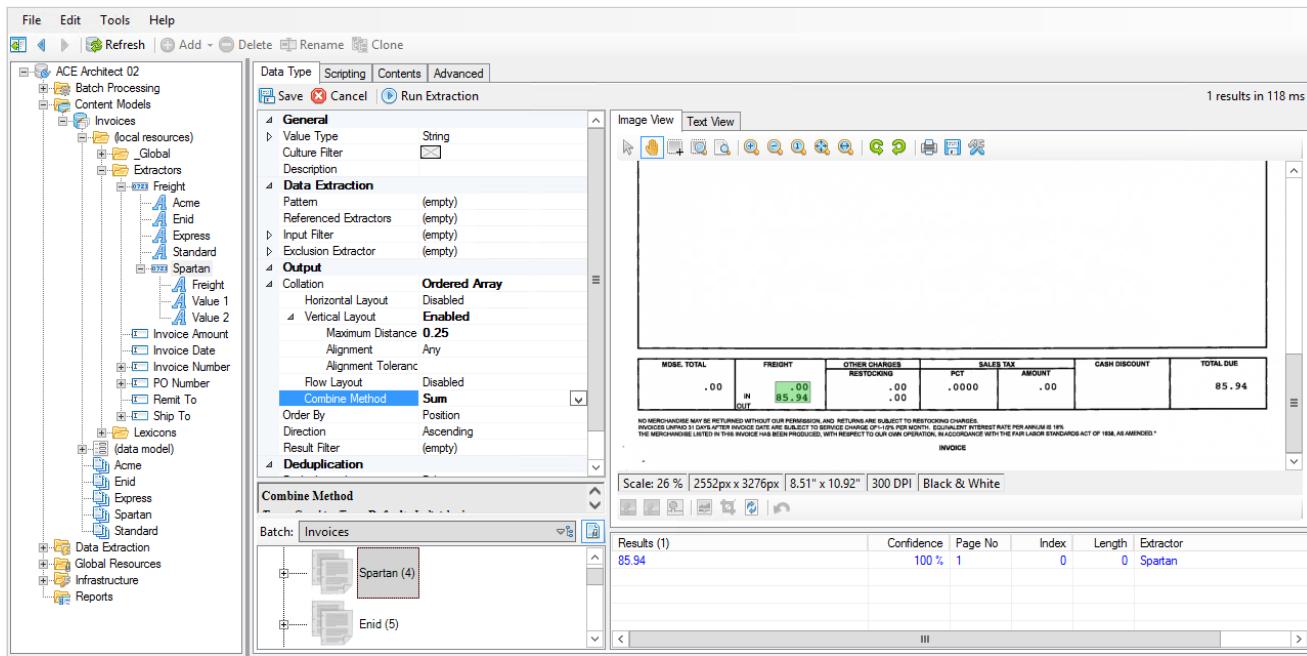
TAKEN BY: DN PAGE 1 OF 1

|                     |                    |                           |                                    |                       |                  |
|---------------------|--------------------|---------------------------|------------------------------------|-----------------------|------------------|
| ORDER DATE 11/10/08 | TERMS 10-30 NET 30 | SHIP DATE 11/10/08        | SHIP VIA DIRECT SHIP & L. CARRIERS | ACCT NUMBER 194306-01 | POB ORG, FRT PPA |
| 10/27/08            | OCN: 123-456-7890  | COMMOD FREIGHT MA-20 ONLY |                                    |                       |                  |

Scale: 26% | 2552px x 3276px | 8.51" x 10.92" | 300 DPI | Black & White

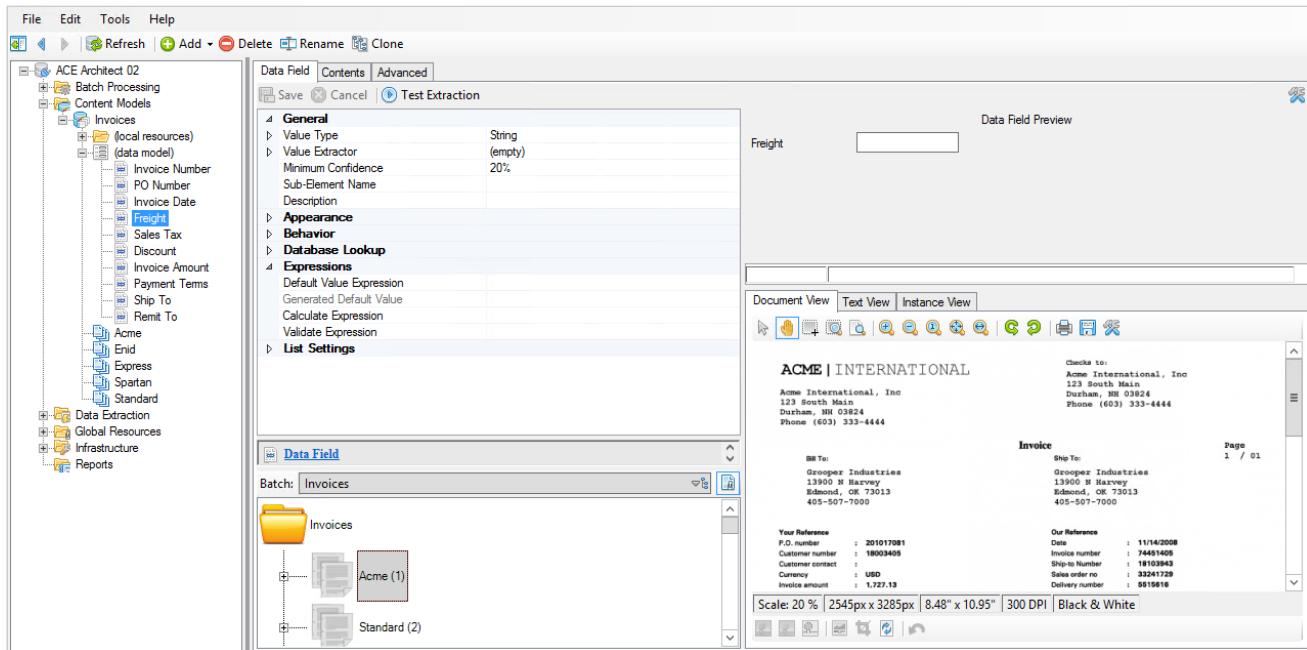
Results (151) Confidence Page No Index Length Extractor

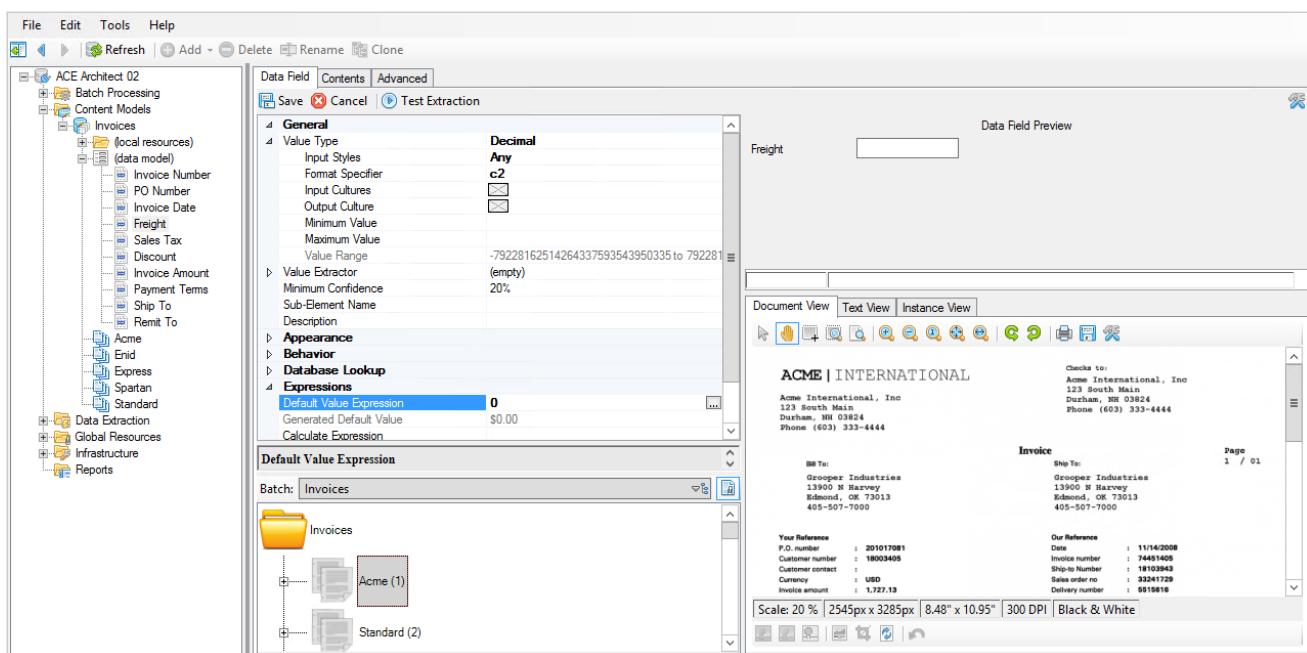
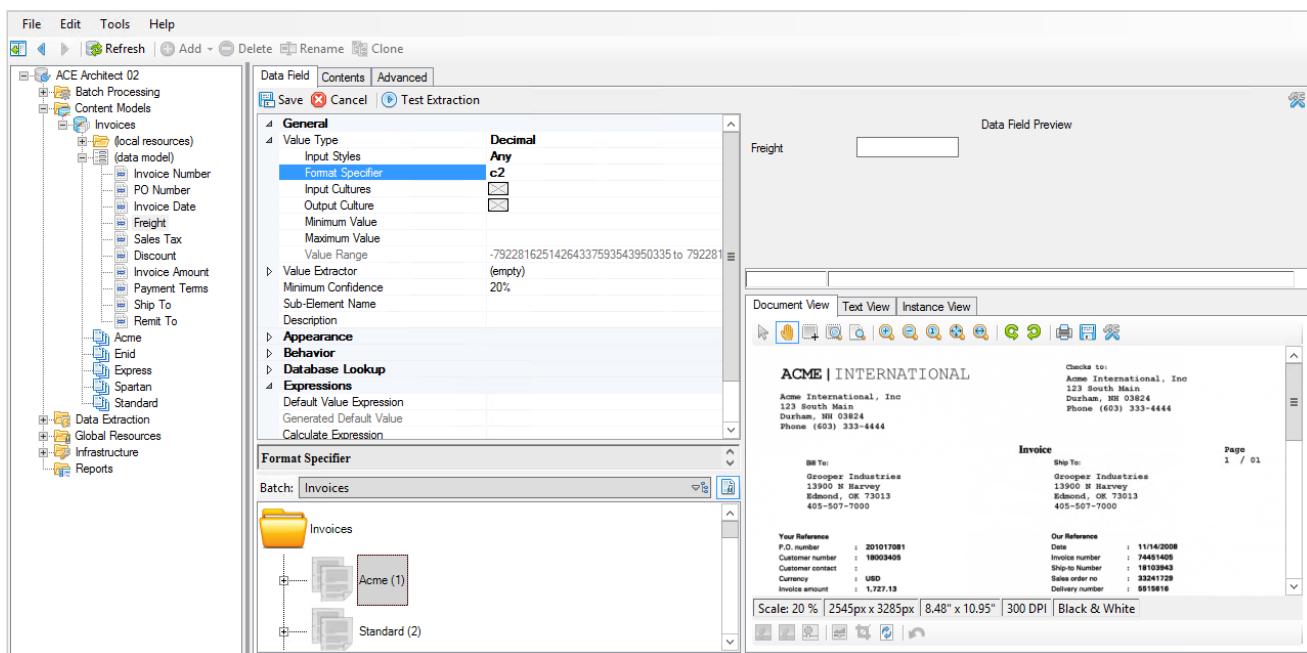
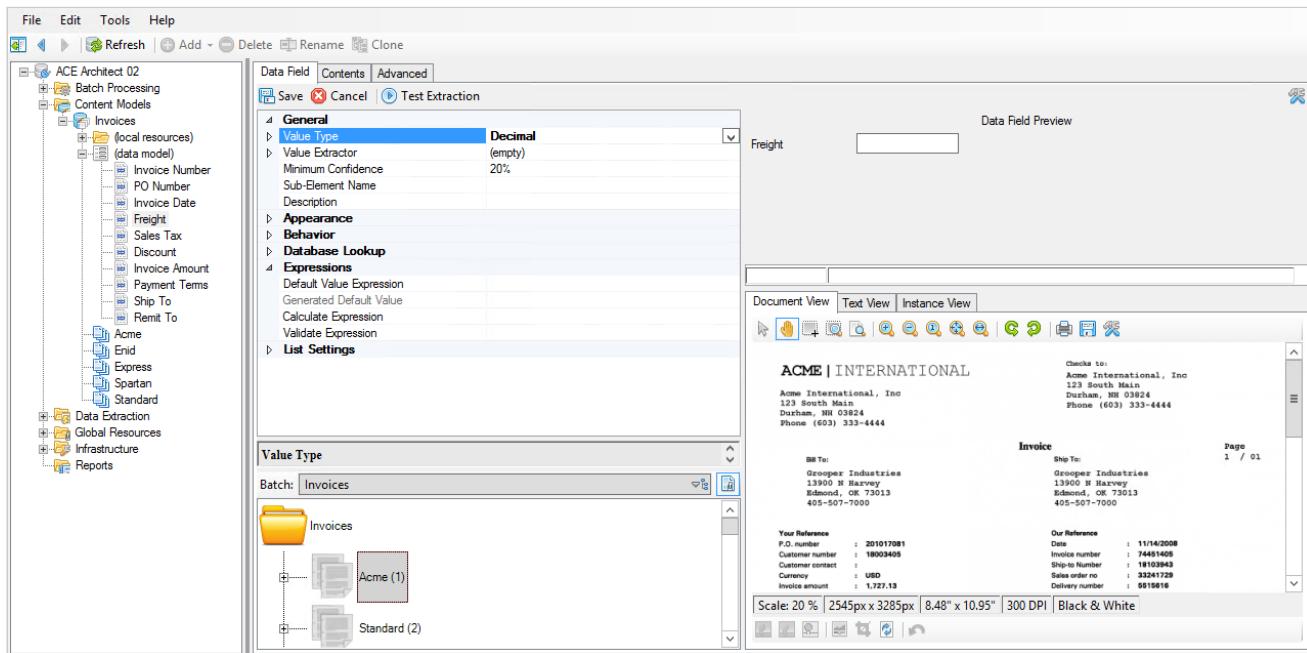
|    |      |   |    |   |         |
|----|------|---|----|---|---------|
| 0  | 100% | 1 | 26 | 1 | Value 1 |
| 0  | 100% | 1 | 26 | 1 | Value 2 |
| 12 | 100% | 1 | 32 | 2 | Value 2 |
| 12 | 100% | 1 | 32 | 2 | Value 1 |
| 0  | 100% | 1 | 63 | 5 | Value 2 |



## Step

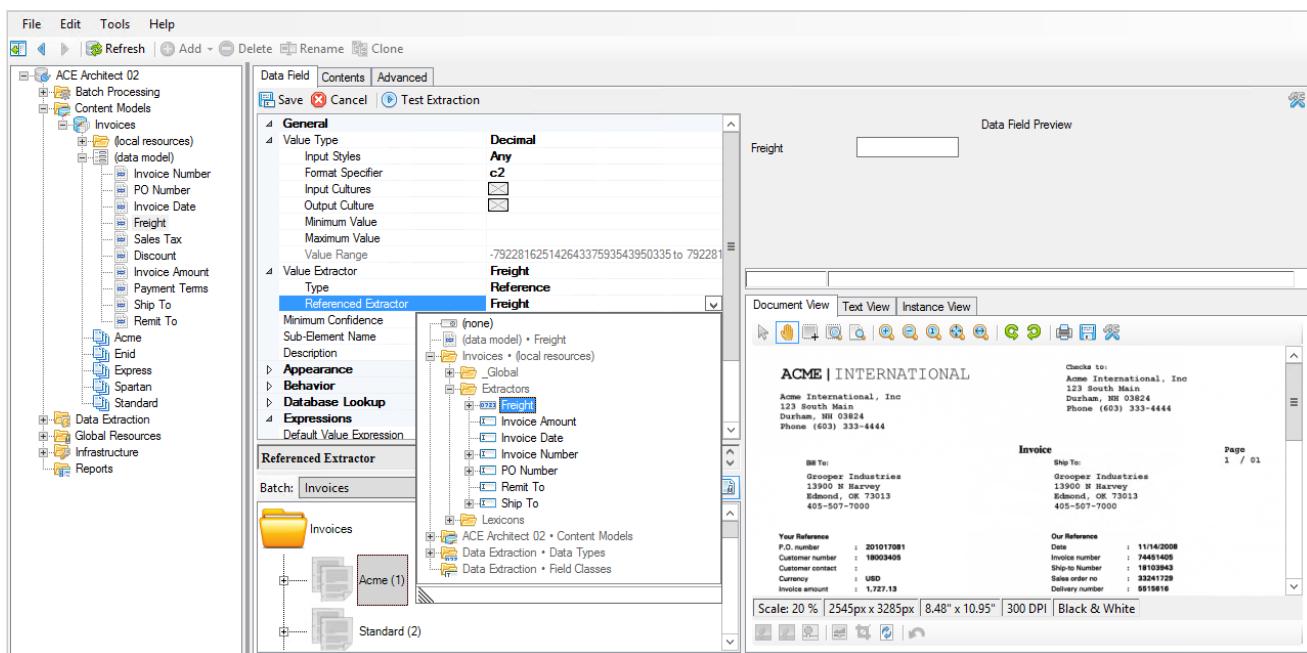
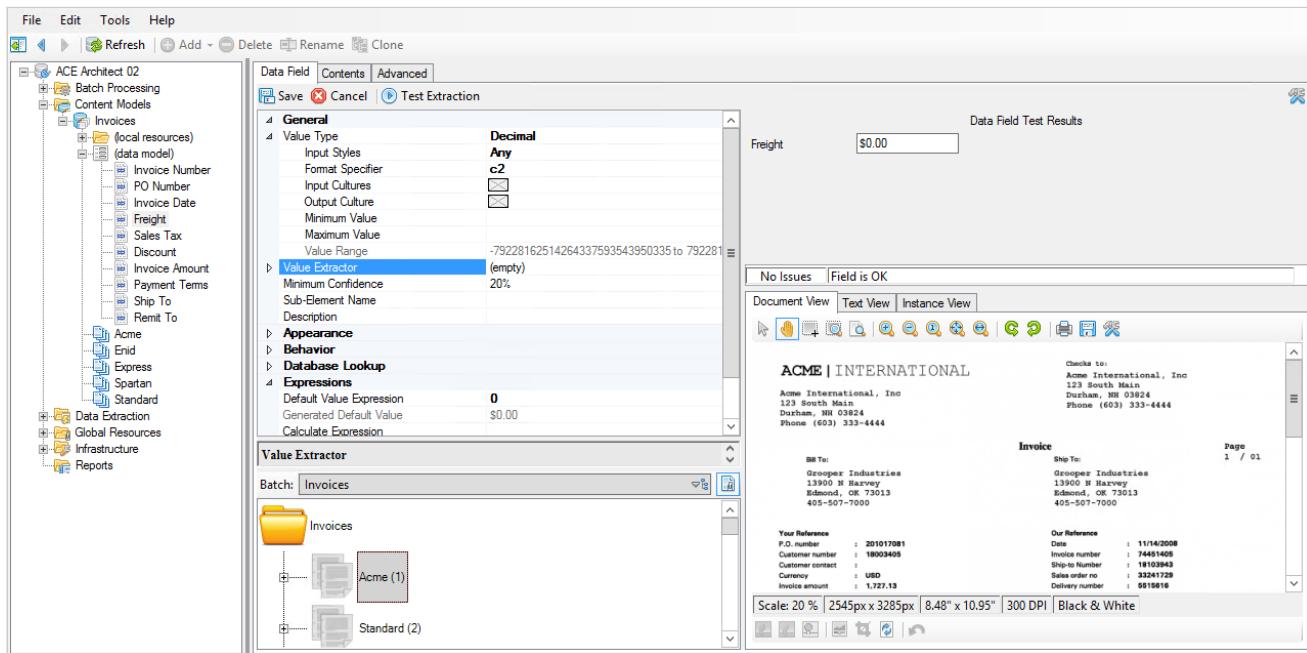
1. In the Data Model, select the Freight Data Field
2. In the General section, set the Value Type to **Decimal**
3. Expand Value Type and set the Format Specifier property to **c2**
4. In the Expressions section, set the Default Value Expression property to **0**





## Step

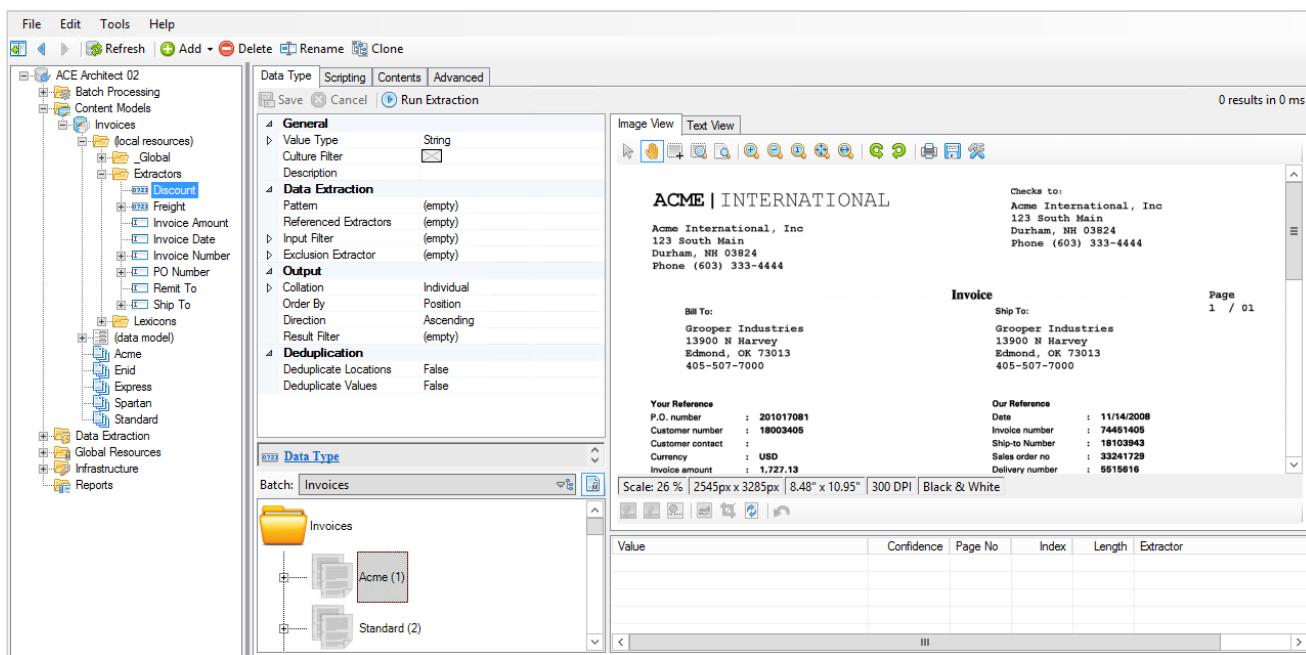
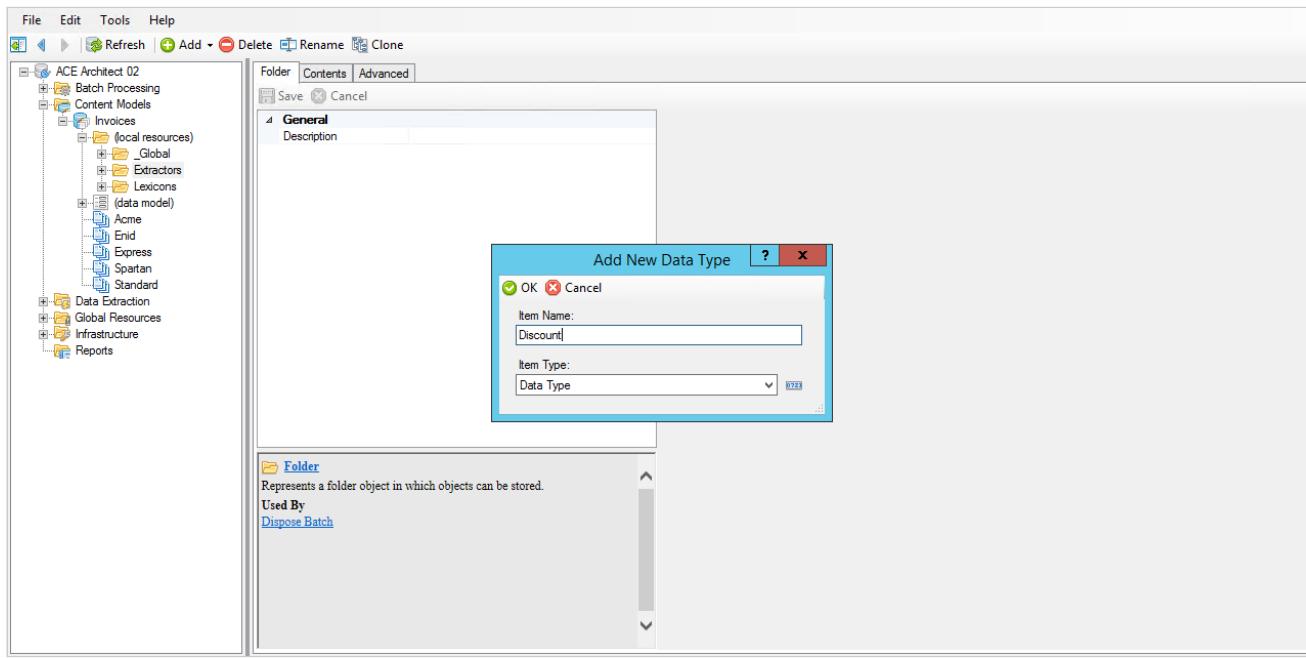
1. In the General section, set the Value Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > Freight Data Type.



## Discount

### Step

1. Right click on the Extractors folder and select Add > Data Type...
2. Name it **Discount**



## Step

1. Create two Data Types as children of the Discount Data Type:

- Acme
- Enid

2. Select the Acme Data Type and click to edit its Pattern

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

Batch Processing Content Models

Invoices (local resources) \_Global Extractors

Extractors

Discount Acme End

Freight Invoice Amount Invoice Date PO Number Remit To Ship To

Lexicons (data model)

Acme End Express Spartan Standard

Data Extraction Global Resources Infrastructure Reports

Data Type Scripting Contents Advanced

Save Cancel Run Extraction

General Value Type String Culture Filter Description

Data Extraction Pattern (empty) Referenced Extractors (empty) Input Filter (empty) Exclusion Extractor (empty)

Output Collation Individual Order By Position Direction Ascending Result Filter (empty)

Deduplication Deduplicate Locations False Deduplicate Values False

Image View Text View

ACME | INTERNATIONAL

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

Invoice 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

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

Scale: 26% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

Data Type

Batch: Invoices

Invoices

Acme (1)

Standard (2)

Value Confidence Page No Index Length Extractor

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

Batch Processing Content Models

Invoices (local resources) \_Global Extractors

Extractors

Discount Acme End

Freight Invoice Amount Invoice Date PO Number Remit To Ship To

Lexicons (data model)

Acme End Express Spartan Standard

Data Extraction Global Resources Infrastructure Reports

Data Type Scripting Contents Advanced

Save Cancel Run Extraction

General Value Type String Culture Filter Description

Data Extraction Pattern (empty) Referenced Extractors (empty) Input Filter (empty) Exclusion Extractor (empty)

Output Collation Individual Order By Position Direction Ascending Result Filter (empty)

Deduplication Deduplicate Locations False Deduplicate Values False

Image View Text View

ACME | INTERNATIONAL

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

Invoice 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

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

Scale: 26% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

Data Type

Batch: Invoices

Invoices

Acme (1)

Standard (2)

Value Confidence Page No Index Length Extractor

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

Batch Processing Content Models

Invoices (local resources) \_Global Extractors

Extractors

Discount Acme End

Freight Invoice Amount Invoice Date PO Number Remit To Ship To

Lexicons (data model)

Acme End Express Spartan Standard

Data Extraction Global Resources Infrastructure Reports

Data Type Scripting Contents Advanced

Save Cancel Run Extraction

General Value Type String Culture Filter Description

Data Extraction Pattern (empty) Referenced Extractors (empty) Input Filter (empty) Exclusion Extractor (empty)

Output Collation Individual Order By Position Direction Ascending Result Filter (empty)

Deduplication Deduplicate Locations False Deduplicate Values False

Image View Text View

ACME | INTERNATIONAL

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

Invoice 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

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

Scale: 26% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

Data Type

Batch: Invoices

Invoices

Acme (1)

Standard (2)

Value Confidence Page No Index Length Extractor

## Step

1. Click on the Properties tab and set the Preprocessing Options > Tab Marking property to **True**

2. Click back to the Pattern Editor tab

3. In the Value Pattern editor, type:

```
[@Number.,]{3,12}
```

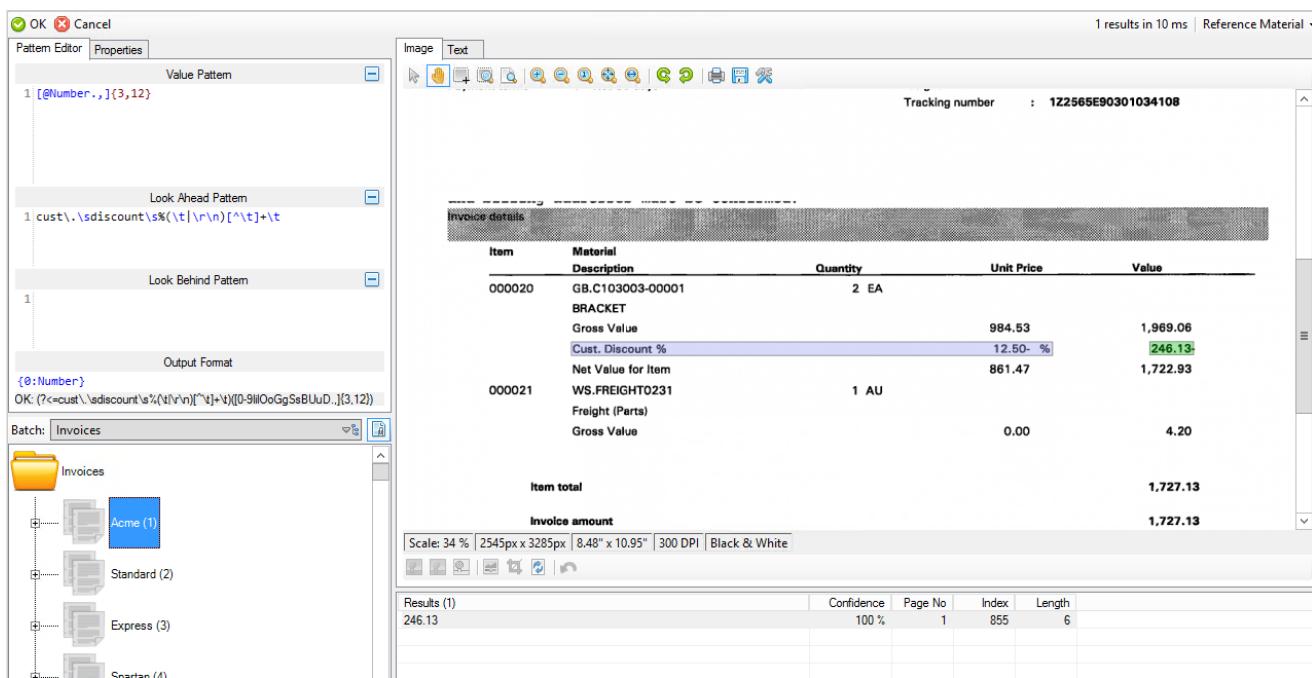
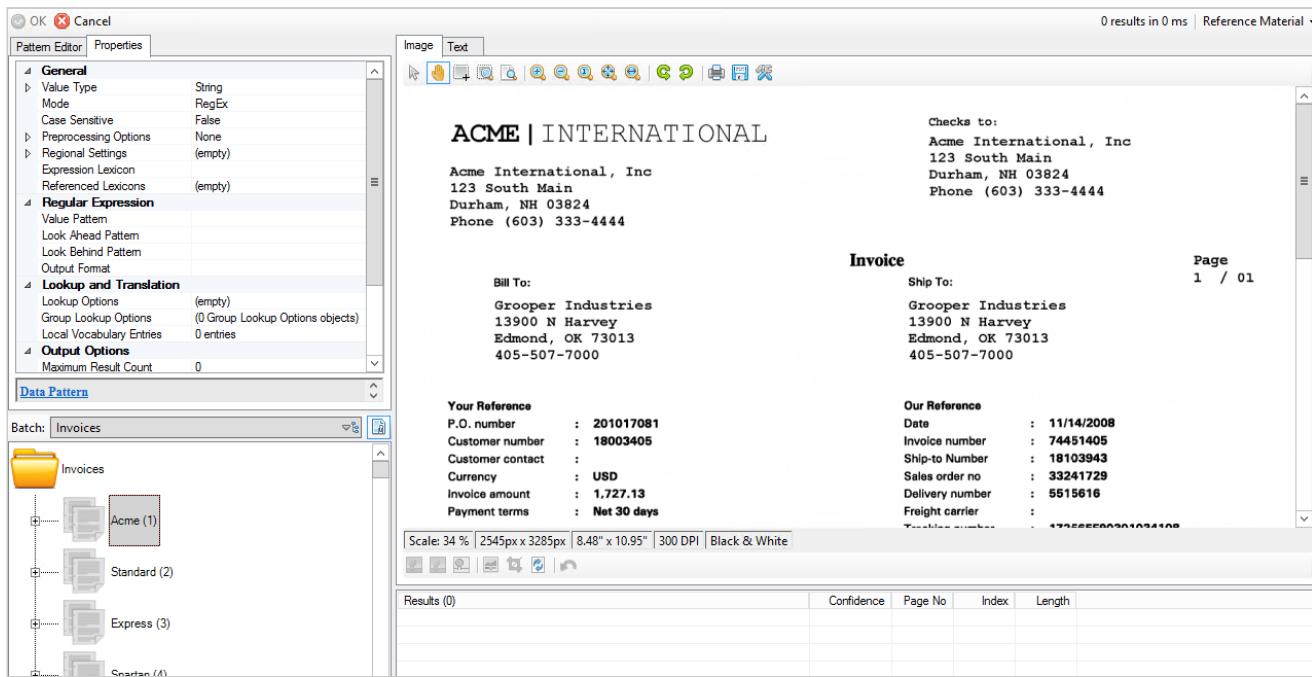
4. In the Look Ahead Pattern editor, type:

```
cust\.\sdiscount\s%(\t|\r\n)[^\t]+\t
```

5. In the Output Format editor, type:

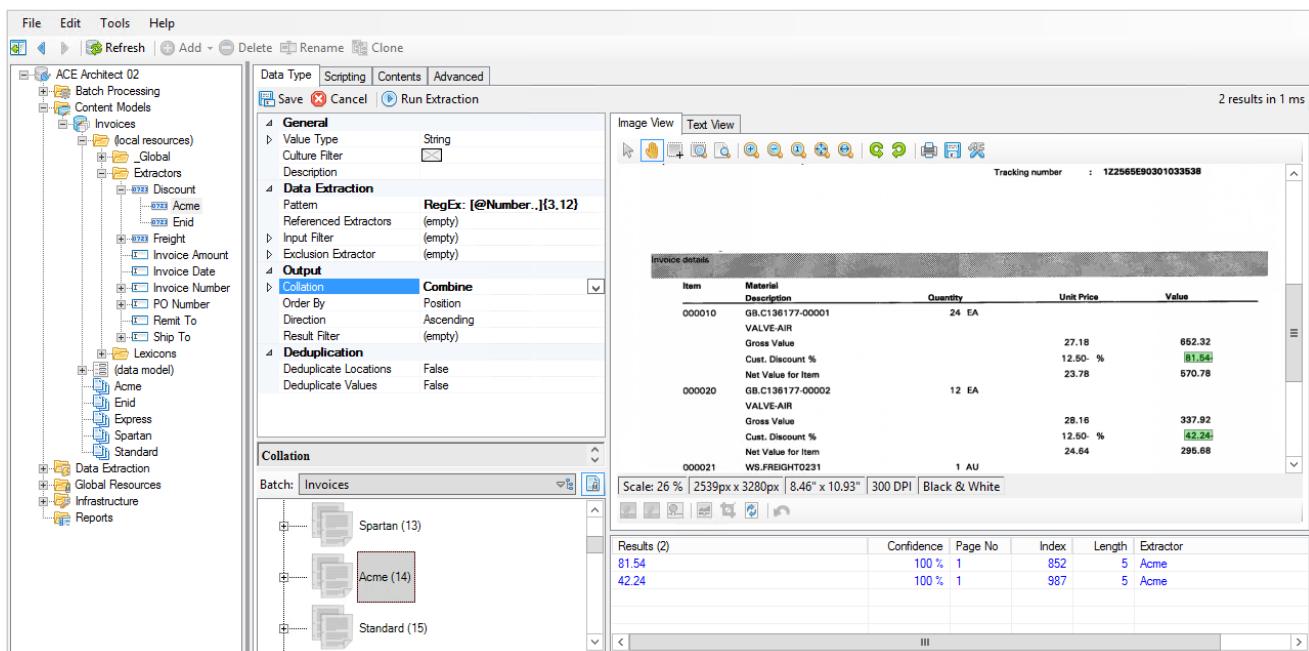
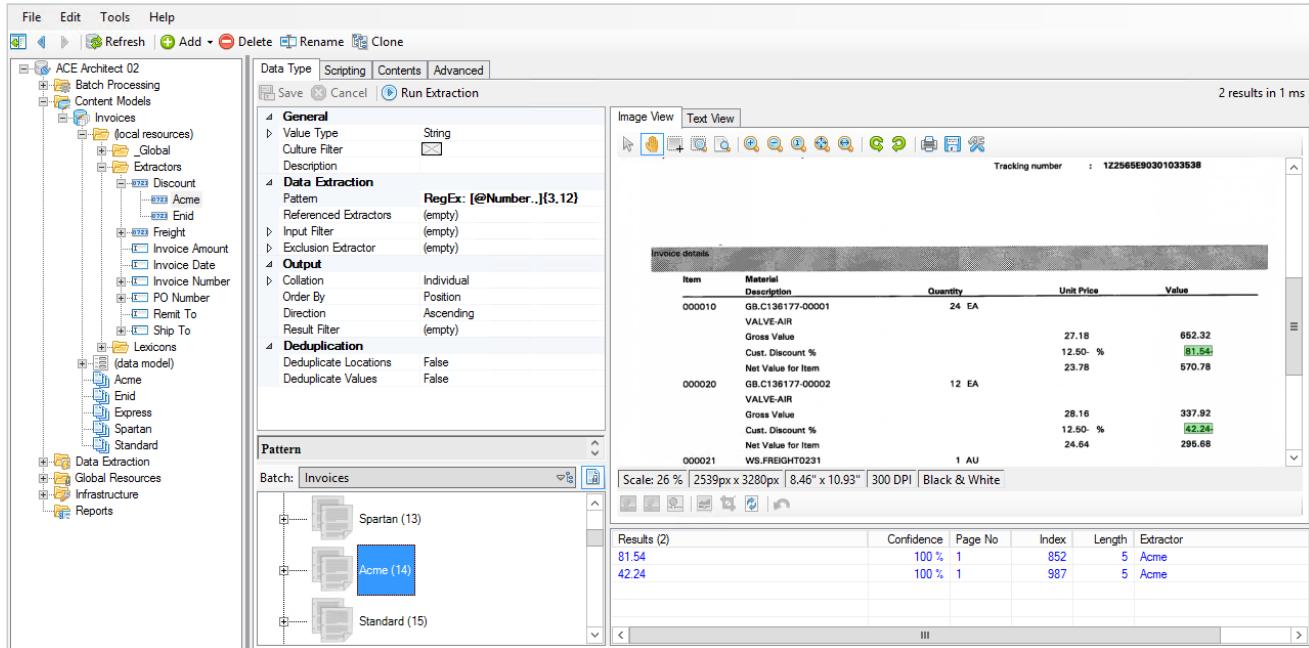
```
{0:Number}
```

6. Click OK to close the pattern editor window



## Step

1. Select Acme (14)
2. In the Output section, set the Collation property to **Combine**
3. Expand Collation and set the Combine Method property to **Sum**

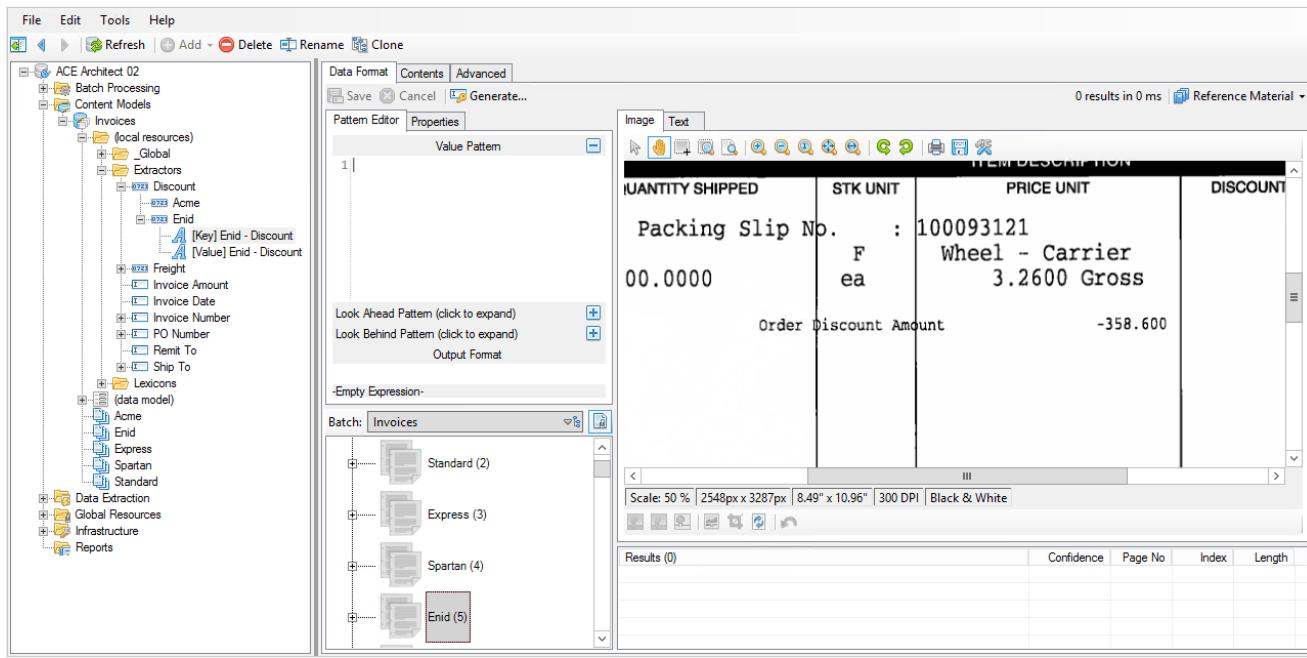


--- Create two Data Formats as children of the Enid Data Type: - [Key] Enid - Discount - [Value] Enid - Discount

## Step

1. Select the [Key] Enid - Discount Data Format
2. Select Enid (5)
3. In the Value Pattern editor, type:

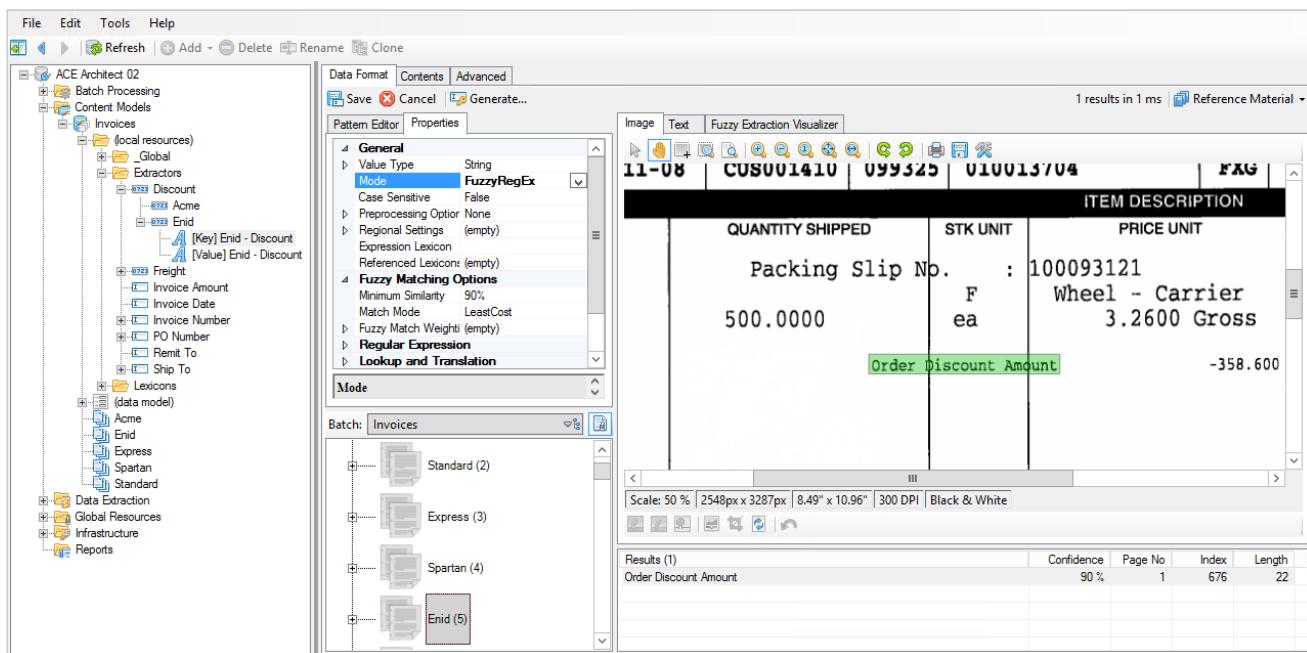
order discount amount

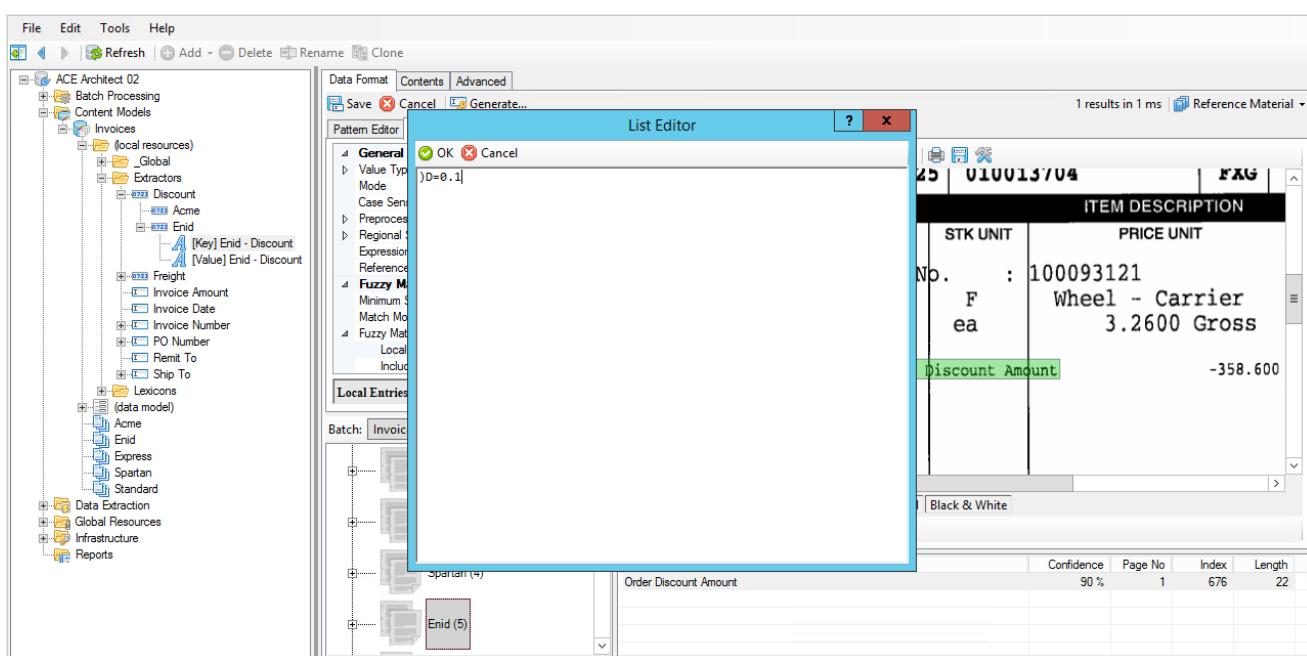
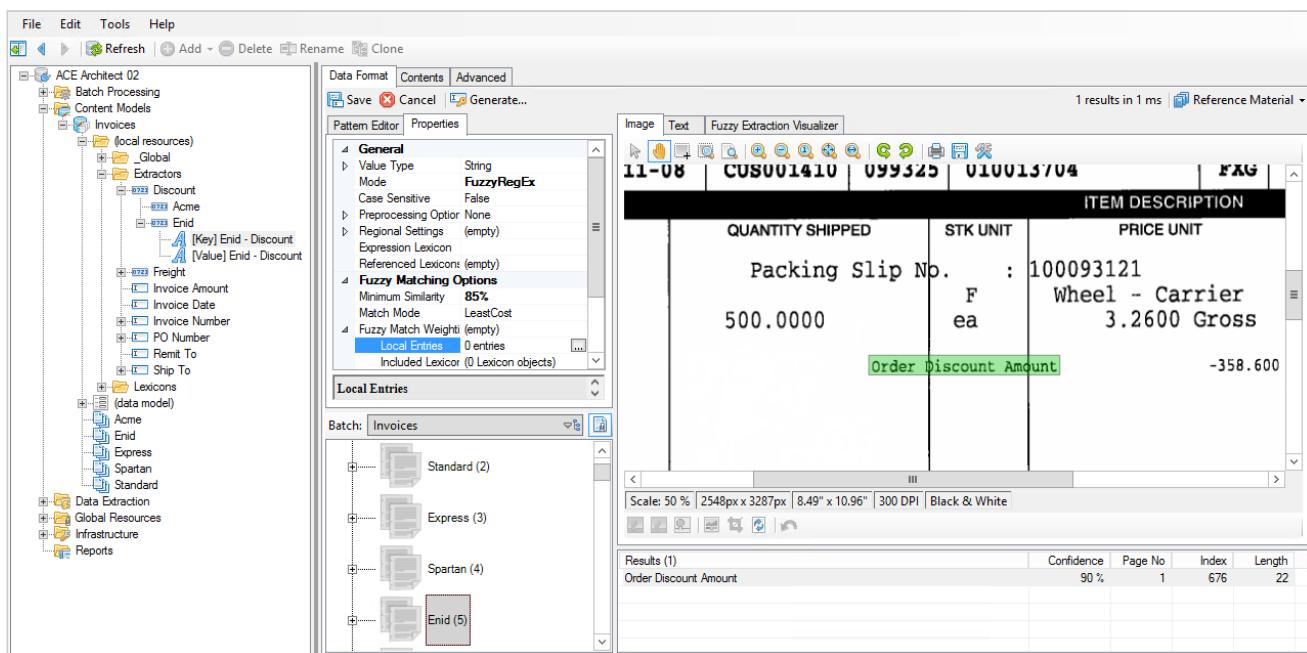
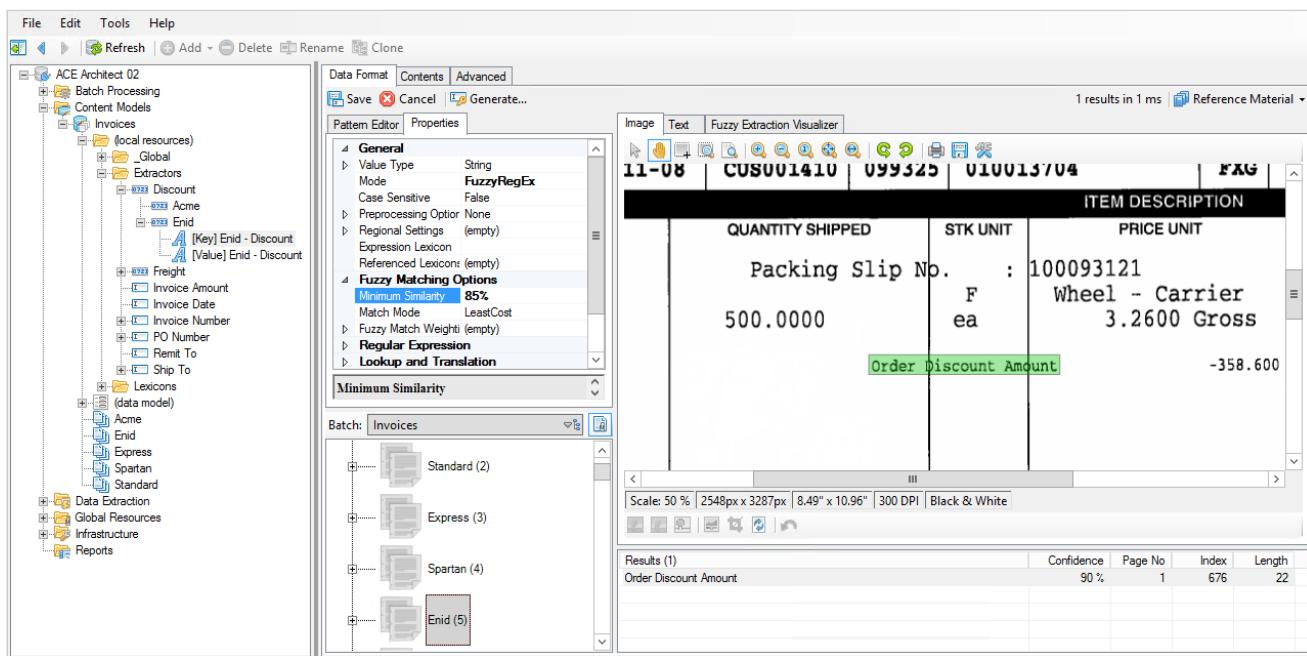


## Step

1. Click on the Properties tab
2. In the General section, set the Mode property to **FuzzyRegEx**
3. In the Fuzzy Matching Options section, set the Minimum Similarity property to **85%**
4. Expand Fuzzy Match Weightings and click to edit Local Entries
5. In the List Editor window, type:  
**)D=0.1**

6. Click OK to close the List Editor window





## Step

1. Select the [Value] Enid - Discount Data Format

2. In the Value Pattern editor, type:

```
\d{1,6}\.\d{2,3}
```

File Edit Tools Help

ACE Architect 02

Batch Processing Content Models Invoices (local resources) Global Extractors Discount Acme Enid [Key] Enid - Discount [Value] Enid - Discount Freight Invoice Amount Invoice Date Invoice Number PO Number Remit To Ship To Lexicons (data model) Acme Enid Express Spartan Standard Data Extraction Global Resources Infrastructure Reports

Data Format Contents Advanced Save Cancel Generate... Pattern Editor Properties

Value Pattern 1

Look Ahead Pattern (click to expand) Look Behind Pattern (click to expand) Output Format

-Empty Expression-

Batch: Invoices

100049766 11-11-08 CUS001410 099325 010013704

PH: (405) 555-1212 FAX: (405) 444-2121

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

SOLD TO: Grocer Industries 13900 N Harvey Edmond, OK 73013

SHIP TO: Grocer Industries 13900 N Harvey Edmond, OK 73013

\*\*\*\*\*

| ORDER NUMBER | ORDER DATE       | CUSTOMER NUMBER  | SALES PERSON     | PURCHASE ORDER NUMBER | SHIP VIA   | SHIP DATE | TERMS          |
|--------------|------------------|------------------|------------------|-----------------------|------------|-----------|----------------|
| 100049766    | 11-11-08         | CUS001410        | 099325           | 010013704             | FXG        | 12-03-08  | Net 30         |
|              |                  |                  |                  |                       |            |           |                |
| ITEM NUMBER  | QUANTITY ORDERED | ITEM DESCRIPTION | QUANTITY SHIPPED | STK UNIT              | PRICE UNIT | DISCOUNT  | EXTENDED PRICE |
|              |                  |                  |                  |                       |            |           |                |

Scale: 25 % | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

Results (0) Confidence Page No Index Length

File Edit Tools Help

ACE Architect 02

Batch Processing Content Models Invoices (local resources) Global Extractors Discount Acme Enid [Key] Enid - Discount [Value] Enid - Discount Freight Invoice Amount Invoice Date Invoice Number PO Number Remit To Ship To Lexicons (data model) Acme Enid Express Spartan Standard Data Extraction Global Resources Infrastructure Reports

Data Format Contents Advanced Save Cancel Generate... Pattern Editor Properties

Value Pattern 1 \d{1,6}\.\d{2,3}

Look Ahead Pattern (click to expand) Look Behind Pattern (click to expand) Output Format

OK:\d{1,6}\.\d{2,3}

Batch: Invoices

100049766 11-11-08 CUS001410 099325 010013704

g Slip No. : 100093121

F ea

Wheel - Carrier

3.2600 Gross

Order Discount Amount -358.600

Results (10)

| Confidence | Page No | Index | Length |
|------------|---------|-------|--------|
| 100 %      | 1       | 602   | 8      |
| 100 %      | 1       | 632   | 7      |
| 100 %      | 1       | 641   | 7      |
| 100 %      | 1       | 653   | 5      |
|            |         |       |        |

Scale: 62 % | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

Results (10) Confidence Page No Index Length

## Step

1. Select the Enid Data Type
2. In the Output section, set the Collation property to **Key-Value Pair**
3. Expand Collation and set the Horizontal Layout property to **Enabled**

File Edit Tools Help

ACE Architect 02

Batch Processing Content Models Invoices (local resources) Global Extractors Discount Acme Enid [Key] Enid - Discount [Value] Enid - Discount Freight Invoice Amount Invoice Date Invoice Number PO Number Remit To Ship To Lexicons (data model) Acme Enid Express Spartan Standard Data Extraction Global Resources Infrastructure Reports

Data Type Scripting Contents Advanced Save Cancel Run Extraction

General Value Type String Culture Filter Description

Data Extraction Pattern (empty) Referenced Extractors (empty) Input Filter (empty) Exclusion Extractor (empty)

Output Collation Individual Order By Position Direction Ascending Result Filter

Deduplication Deduplicate Locations False Deduplicate Values False

Data Type Batch: Invoices

100049766 11-11-08 CUS001410 099325 010013704

\*\*\*\*\*

| ORDER NUMBER | ORDER DATE       | CUSTOMER NUMBER  | SALES PERSON     | PURCHASE ORDER NUMBER | SHIP VIA   | SHIP DATE | TERMS          |
|--------------|------------------|------------------|------------------|-----------------------|------------|-----------|----------------|
| 100049766    | 11-11-08         | CUS001410        | 099325           | 010013704             | FXG        | 12-03-08  | Net 30         |
|              |                  |                  |                  |                       |            |           |                |
| ITEM NUMBER  | QUANTITY ORDERED | ITEM DESCRIPTION | QUANTITY SHIPPED | STK UNIT              | PRICE UNIT | DISCOUNT  | EXTENDED PRICE |
|              |                  |                  |                  |                       |            |           |                |

7706.105 : 100093121

500.000 F ea

500.000 Wheel - Carrier

3.2600 Gross

Order Discount Amount -358.600

Results (11)

| Confidence | Page No | Index | Length |
|------------|---------|-------|--------|
| 100 %      | 1       | 602   | 8      |
| 100 %      | 1       | 632   | 7      |
| 100 %      | 1       | 641   | 7      |
| 100 %      | 1       | 653   | 5      |
|            |         |       |        |

Scale: 24 % | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

Results (11) Confidence Page No Index Length Extractor

## Step

1. In the Data Model, select the Discount Data Field
2. In the General section, set the Value Type to **Decimal**
3. Expand Value Type and set the Format Specifier property to **c2**
4. In the Expressions section, set the Default Value Expression property to **0**

**ACE Architect 02**

- Batch Processing
- Content Models
  - Invoices
    - (local resources)
    - (data model)
    - Invoice Number
    - PO Number
    - Invoice Date
    - Freight
    - Sales Tax
    - Discount**
    - Invoice Amount
    - Payment Terms
    - Ship To
    - Remit To
- Acme
- End
- Express
- Spartan
- Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

**Data Field** **Contents** **Advanced**

**Save** **Cancel** **Test Extraction**

**General**

- Value Type: String
- Value Extractor: (empty)
- Minimum Confidence: 20%
- Sub-Element Name:
- Description:

**Appearance**

**Behavior**

**Database Lookup**

**Expressions**

- Default Value Expression
- Generated Default Value
- Calculate Expression
- Validate Expression

**List Settings**

**Data Field**

Batch: Invoices

**Invoices**

- Acme (1)
- Standard (2)

**Discount**

**Data Field Preview**

Document View Text View Instance View

**ACME | INTERNATIONAL**

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

**Invoice**

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

Page 1 / 01

**Reference**

|                  |             |                 |                   |
|------------------|-------------|-----------------|-------------------|
| P.O. number      | 201017081   | Date            | 11/14/2008        |
| Customer number  | 18003405    | Invoice number  | 74451405          |
| Customer contact |             | Ship-to Number  | 18102943          |
| Currency         | USD         | Sales order no. | 33241729          |
| Invoice amount   | 1,227.13    | Delivery method | 5515016           |
| Payment terms    | Net 30 days | Freight carrier |                   |
|                  |             | Tracking number | 123568E0301034108 |

Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

**ACE Architect 02**

- Batch Processing
- Content Models
  - Invoices
    - (local resources)
    - (data model)
    - Invoice Number
    - PO Number
    - Invoice Date
    - Freight
    - Sales Tax
    - Discount**
    - Invoice Amount
    - Payment Terms
    - Ship To
    - Remit To
- Acme
- End
- Express
- Spartan
- Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

**Data Field** **Contents** **Advanced**

**Save** **X Cancel** **Test Extraction**

**General**

- Value Type: Decimal
- Input Styles: Any
- Format Specifier: c2
- Input Cultures:
- Output Culture:
- Minimum Value:
- Maximum Value:
- Value Range: -79228162514264337593543950335 to 792
- Value Extractor: (empty)
- Minimum Confidence: 20%
- Sub-Element Name:
- Description:

**Appearance**

**Behavior**

**Database Lookup**

**Expressions**

- Default Value Expression
- Generated Default Value
- Calculate Expression

**Validate Expression**

Batch: Invoices

**Invoices**

- Acme (1)
- Standard (2)

**Discount**

**Data Field Preview**

Document View Text View Instance View

**ACME | INTERNATIONAL**

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

**Invoice**

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

Page 1 / 01

**Reference**

|                  |             |                 |                   |
|------------------|-------------|-----------------|-------------------|
| P.O. number      | 201017081   | Date            | 11/14/2008        |
| Customer number  | 18003405    | Invoice number  | 74451405          |
| Customer contact |             | Ship-to Number  | 18102943          |
| Currency         | USD         | Sales order no. | 33241729          |
| Invoice amount   | 1,227.13    | Delivery method | 5515016           |
| Payment terms    | Net 30 days | Freight carrier |                   |
|                  |             | Tracking number | 123568E0301034108 |

Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

**ACE Architect 02**

- Batch Processing
- Content Models
  - Invoices
    - (local resources)
    - (data model)
    - Invoice Number
    - PO Number
    - Invoice Date
    - Freight
    - Sales Tax
    - Discount**
    - Invoice Amount
    - Payment Terms
    - Ship To
    - Remit To
- Acme
- End
- Express
- Spartan
- Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

**Data Field** **Contents** **Advanced**

**Save** **X Cancel** **Test Extraction**

**General**

- Value Type: Decimal
- Input Styles: Any
- Format Specifier: c2
- Input Cultures:
- Output Culture:
- Minimum Value:
- Maximum Value:
- Value Range: -79228162514264337593543950335 to 792
- Value Extractor: (empty)
- Minimum Confidence: 20%
- Sub-Element Name:
- Description:

**Expressions**

- Default Value Expression: 0
- Generated Default Value: \$0.00
- Calculate Expression:

**Default Value Expression**

Batch: Invoices

**Invoices**

- Acme (1)
- Standard (2)

**Discount**

**Data Field Preview**

Document View Text View Instance View

**ACME | INTERNATIONAL**

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

**Invoice**

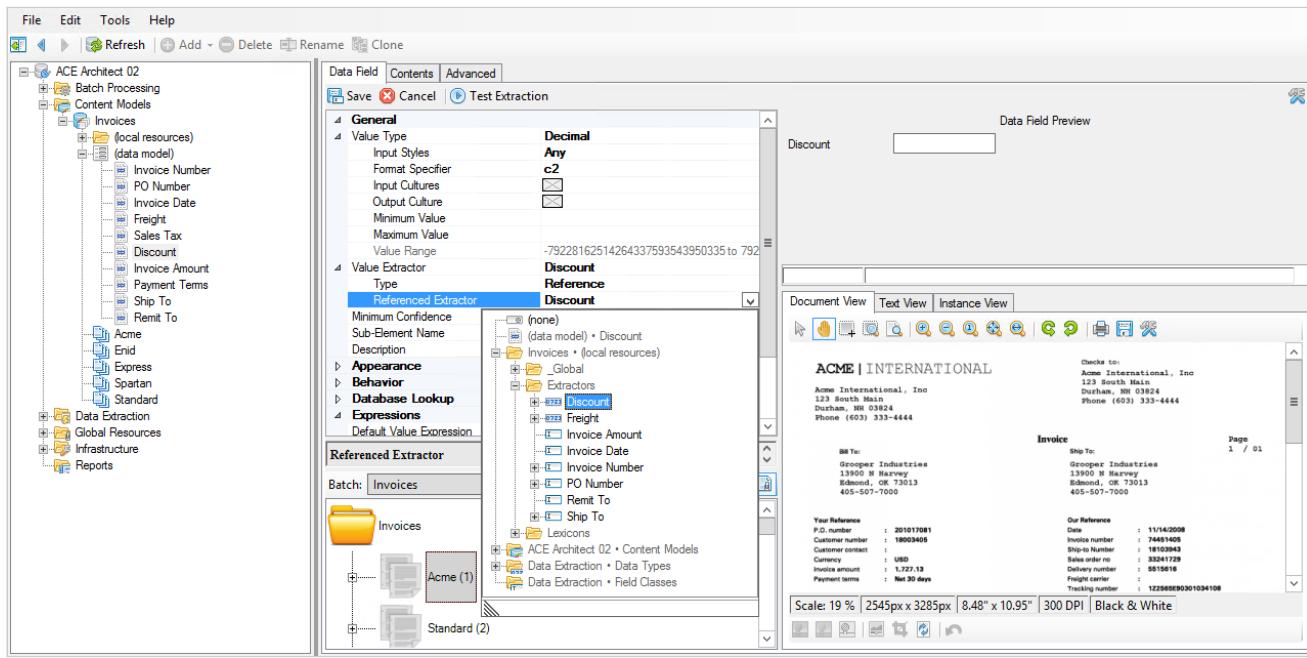
Ship To: Grocoer Industries  
13900 N Harvey  
Edmond, OK 73013  
405-507-7000

Page 1 / 01

**Reference**

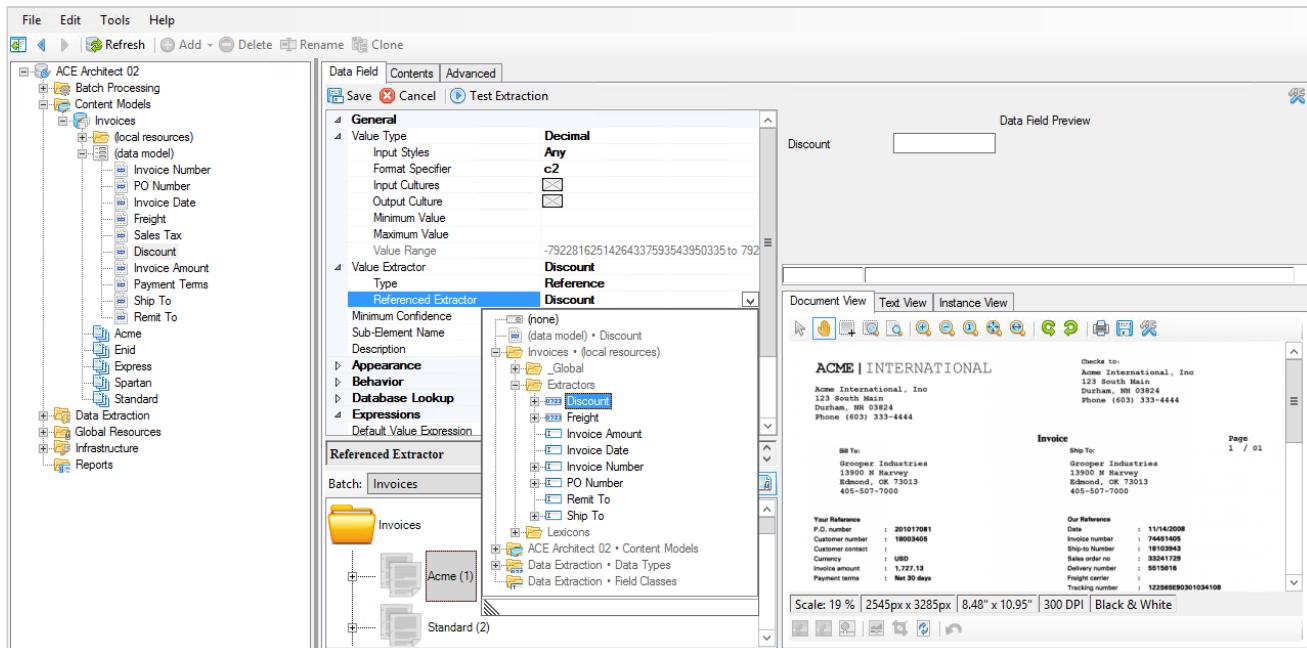
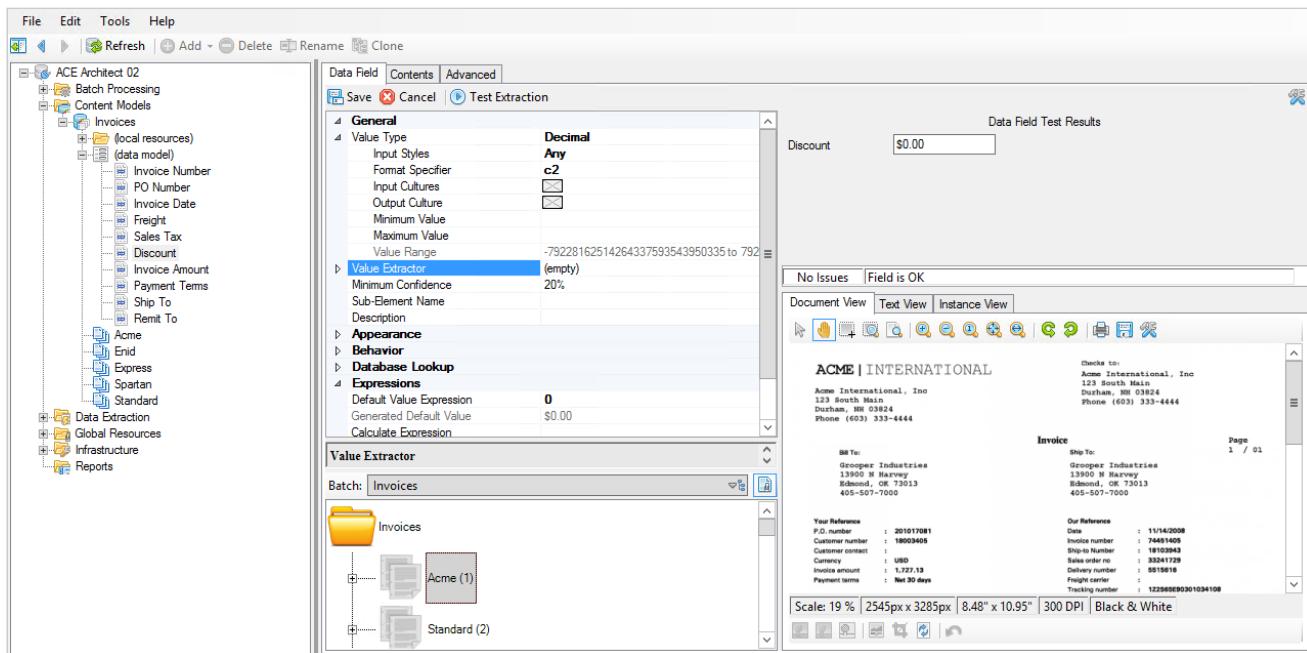
|                  |             |                 |                   |
|------------------|-------------|-----------------|-------------------|
| P.O. number      | 201017081   | Date            | 11/14/2008        |
| Customer number  | 18003405    | Invoice number  | 74451405          |
| Customer contact |             | Ship-to Number  | 18102943          |
| Currency         | USD         | Sales order no. | 33241729          |
| Invoice amount   | 1,227.13    | Delivery method | 5515016           |
| Payment terms    | Net 30 days | Freight carrier |                   |
|                  |             | Tracking number | 123568E0301034108 |

Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White



## Step

1. In the General section, set the Value Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > **Discount** Data Type.



# Sales Tax

## Step

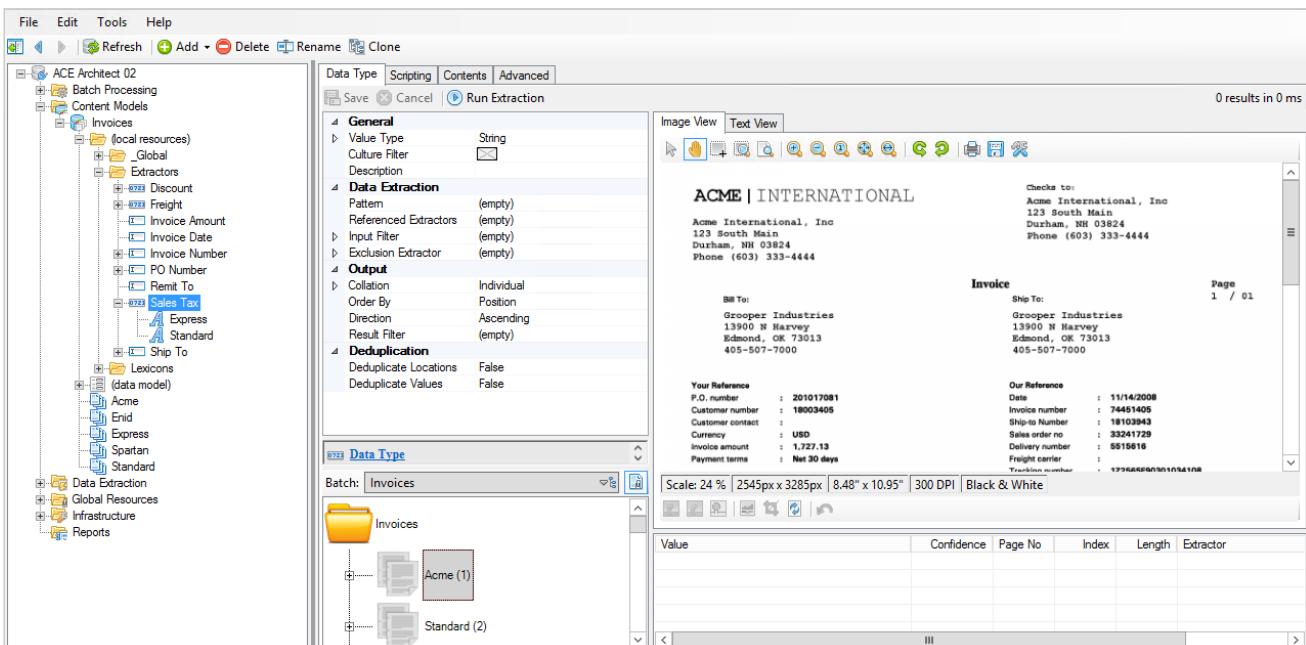
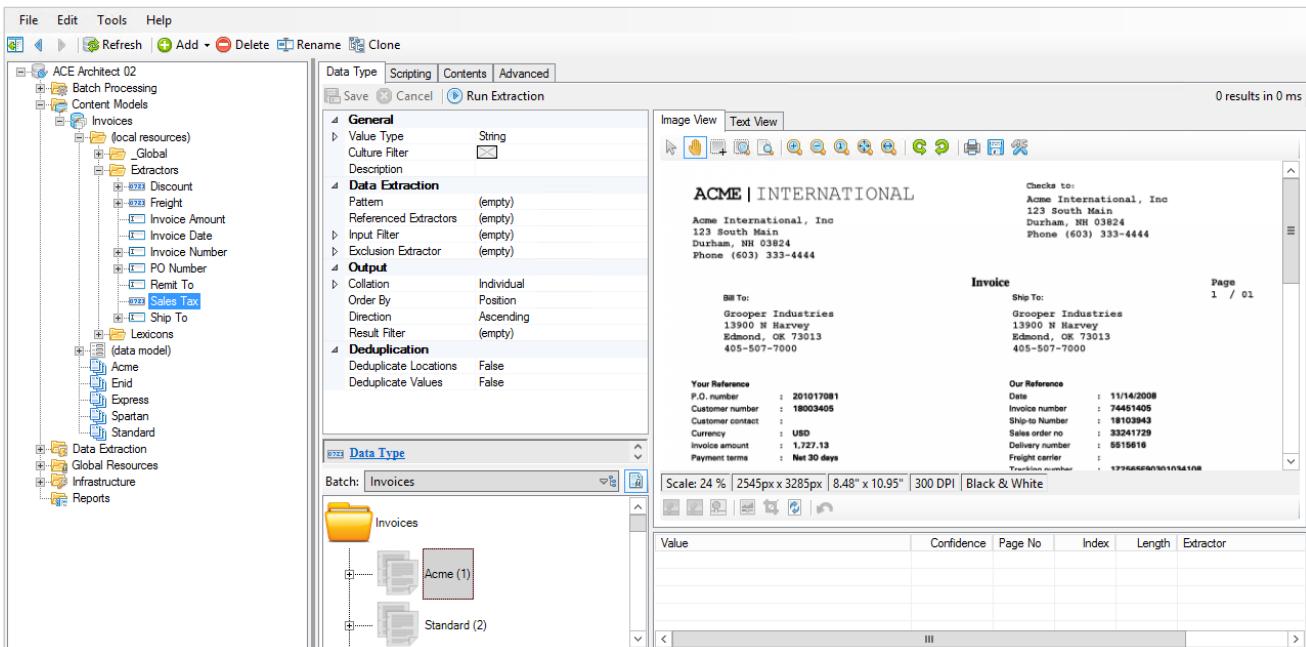
1. Right click on the Extractors folder and select Add > Data Type...

2. Name it **Sales Tax**

3. Create two Data Formats as children of the Sales Tax Data Type

- Express

- Standard



--- Edit the patterns for each of the Data Formats: - Express

```
```tab="Value Pattern"
([@Number.] {4,12})
```

```

```
```tab="Look Ahead Pattern"
sales tax:(\r\n|\s)
```

```
```tab="Output Format"
```

```
{0:Number}
```

```
```
```

- Standard

```
```tab="Value Pattern"
```

```
([@Number.]{3,12})
```

```
```
```

```
```tab="Look Ahead Pattern"
```

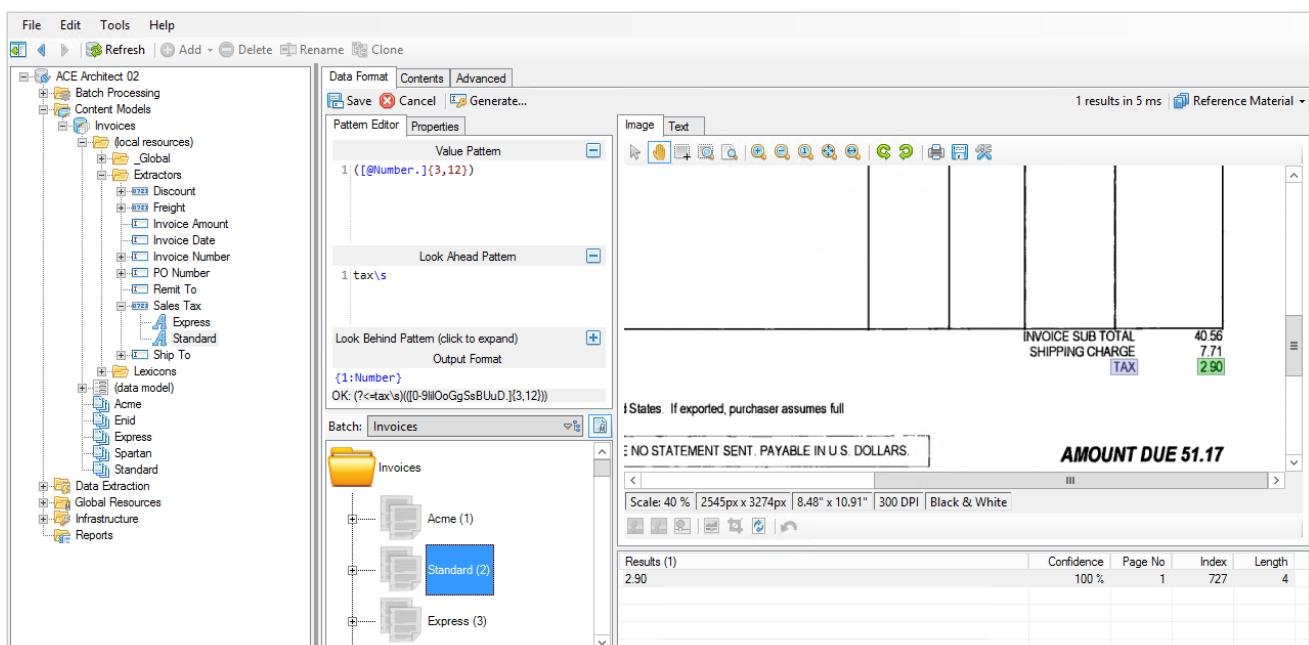
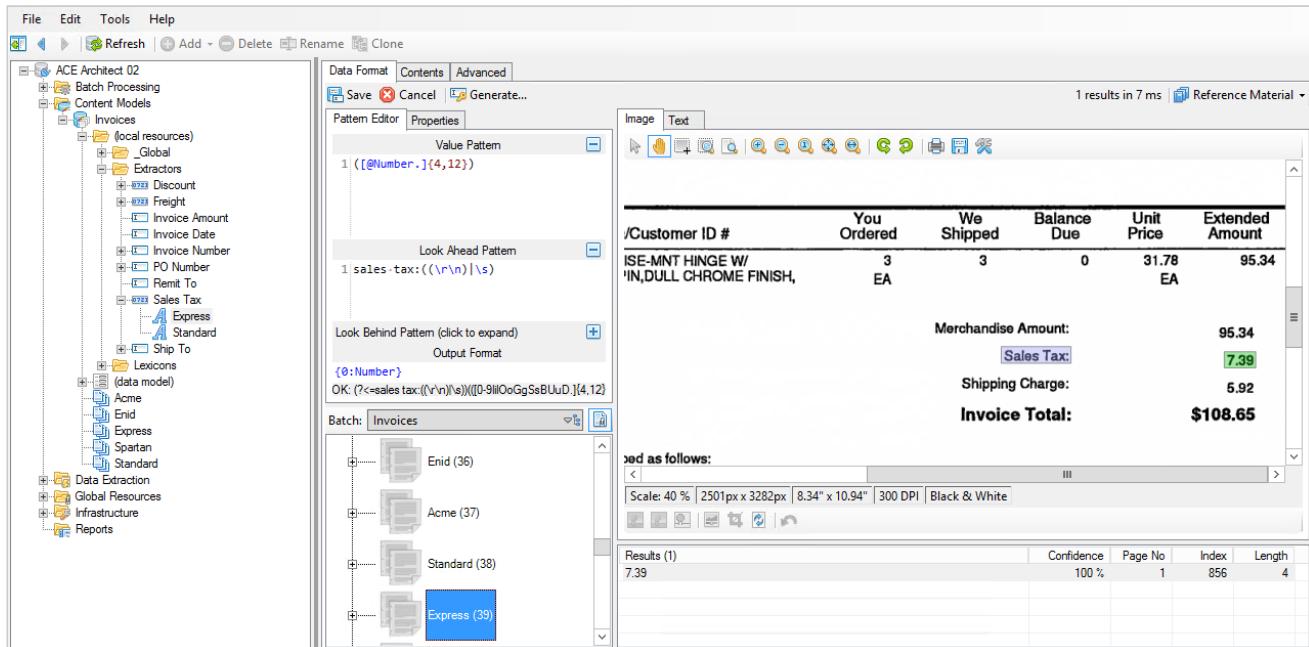
```
tax\s
```

```
```
```

```
```tab="Output Format"
```

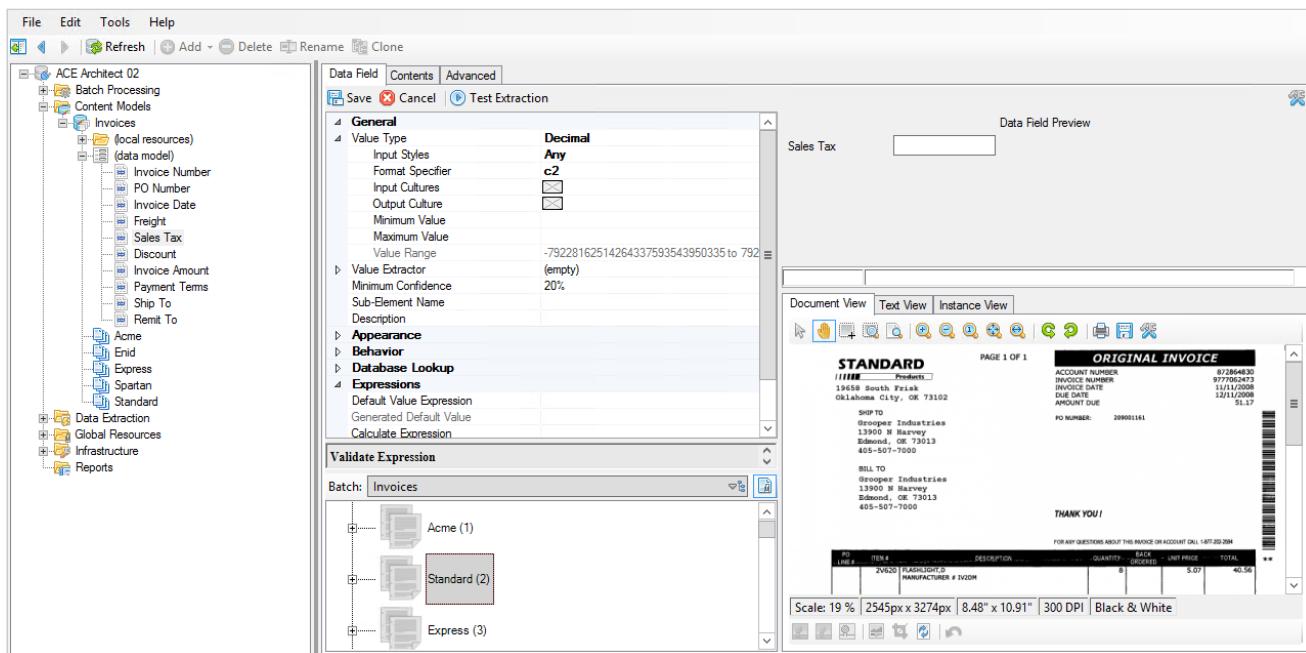
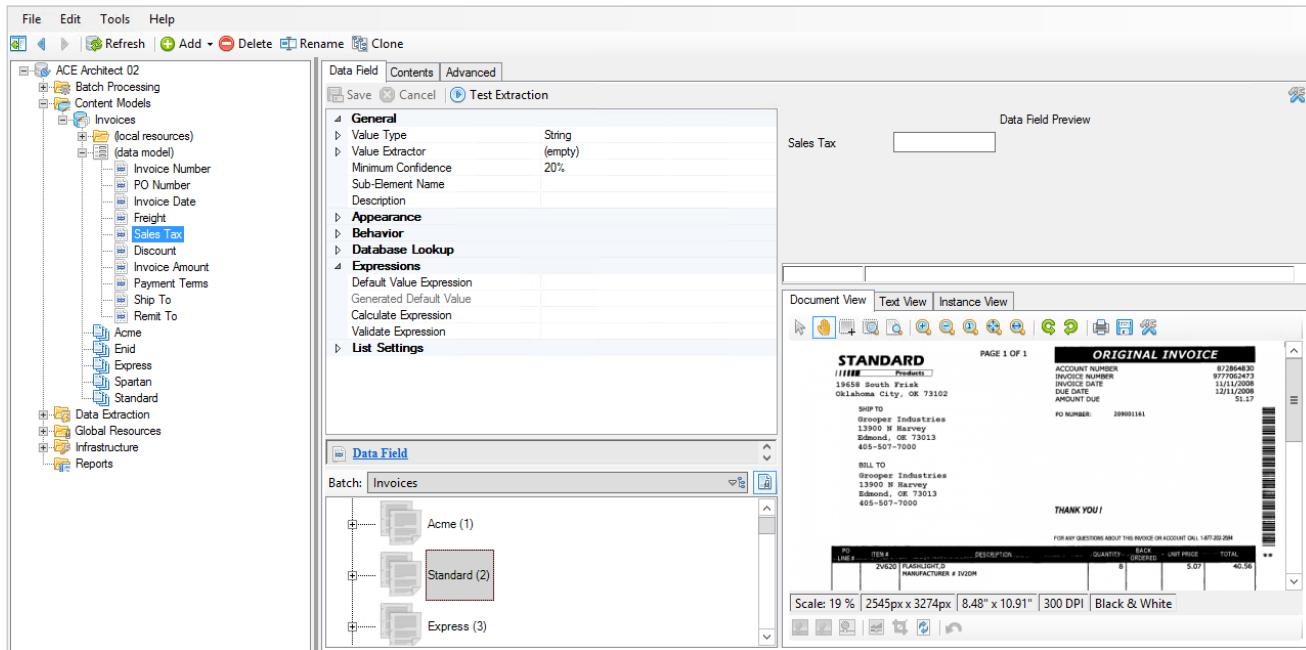
```
{1:Number}
```

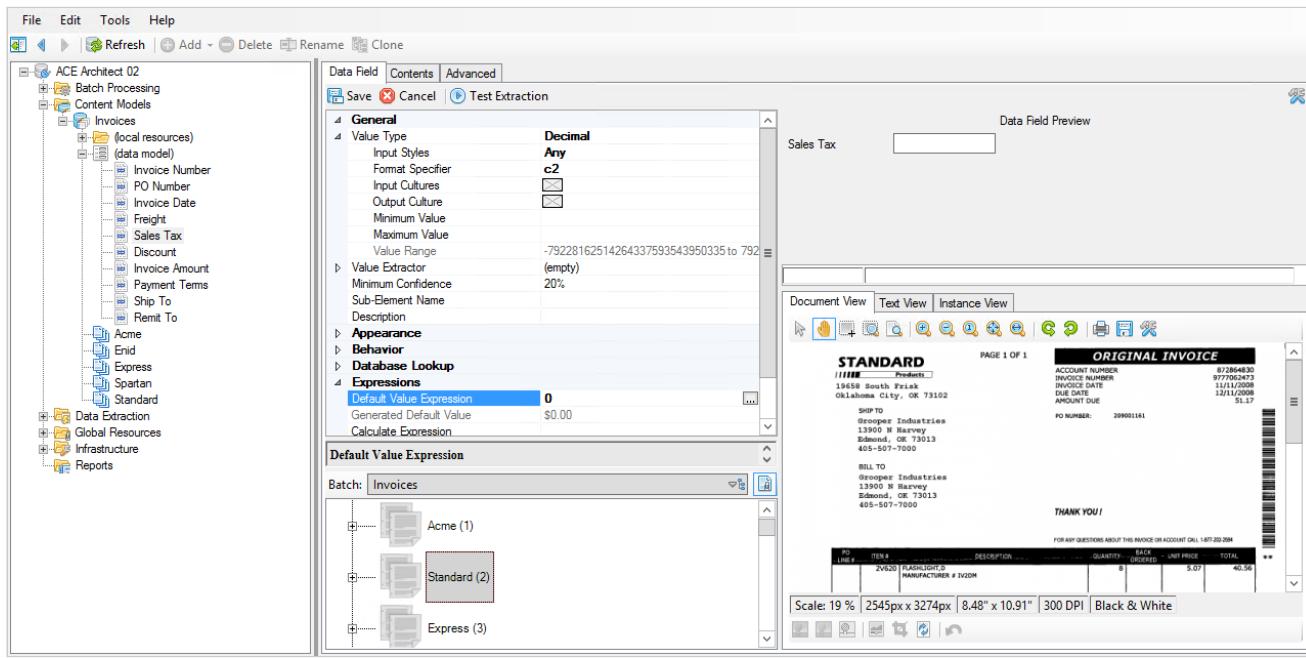
```
```
```



Step

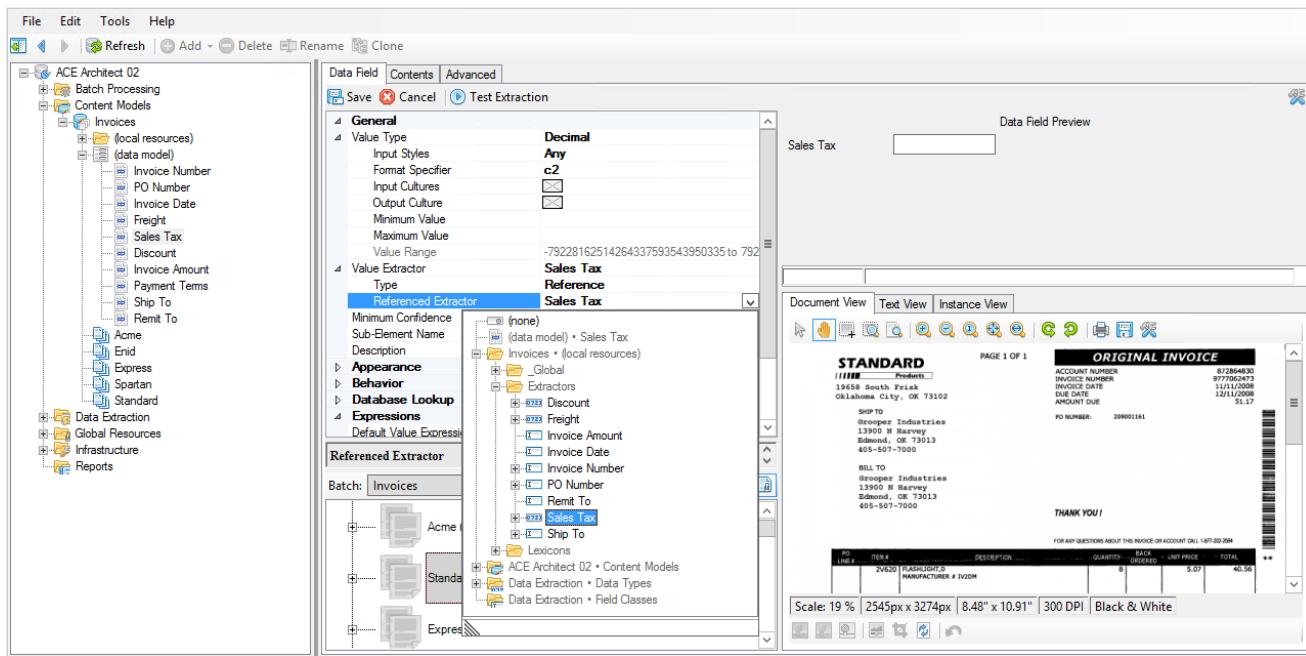
1. In the Data Model, select the Sales Tax Data Field
2. In the General section, set the Value Type to **Decimal**
3. Expand Value Type and set the Format Specifier property to **c2**
4. In the Expressions section, set the Default Value Expression property to **0**





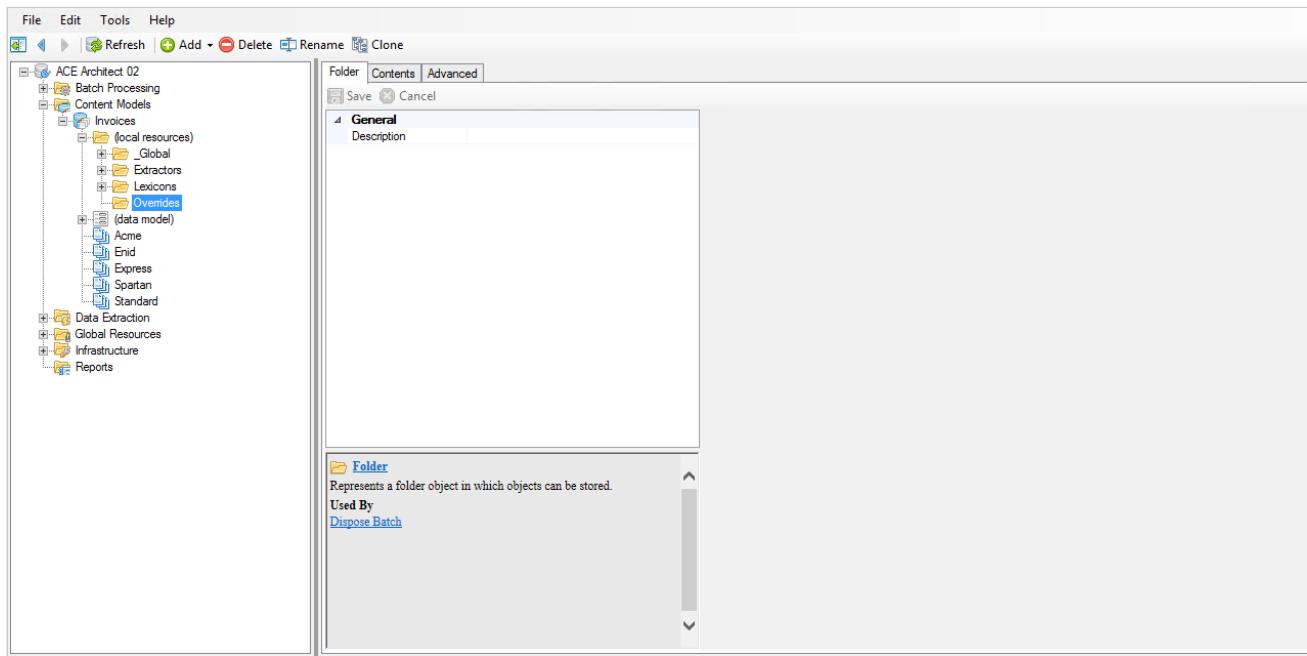
Step

1. In the General section, set the Value Extractor > Type property to Reference
2. Set the Referenced Extractor property to the Invoices • (local resources) > Extractors > **Sales Tax** Data Type.



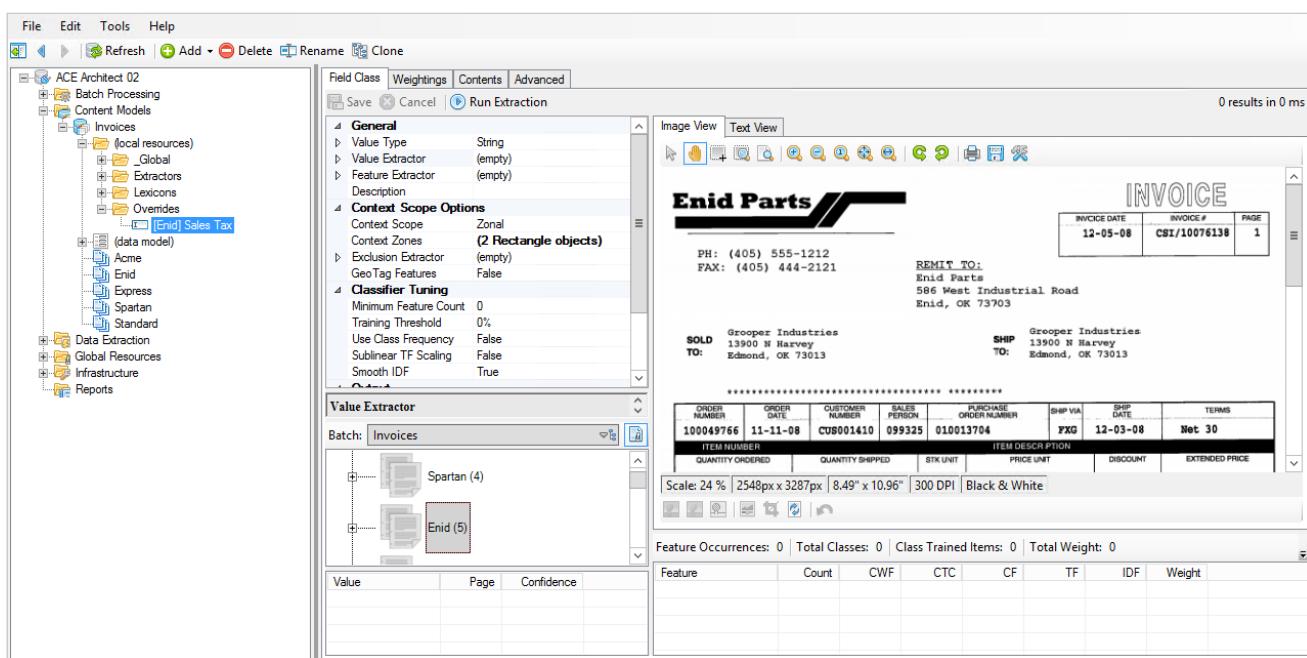
Step

1. Add a new folder to (local resources)
2. Name it **Overrides**



Step

1. In the Overrides folder, add a Field Class
2. Name it [Enid] Sales Tax



Step

1. Set the Value Extractor > Type property to Internal
2. Click to edit the pattern:

Value Pattern

([@Number.] {2,12})

Output Format

3. Set the Feature Extractor > Type property to Reference
4. Set the Referenced Extractor property to the Data Extraction • Data Types > Training Materials > Features > Phrases Data Type

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - _Global
 - Extractors
 - Lexicons
 - Ovemedes
 - [Enid] Sales Tax (data model)
 - Acme
 - Enid
 - Express
 - Spartan
 - Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Field Class Weightings Contents Advanced

Save Cancel Run Extraction

Image View Text View

Enid Parts

INVOICE

INVOICE DATE: 12-05-08 INVOICE #: CSI/10076138 PAGE: 1

PH: (405) 555-1212 FAX: (405) 444-2121

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

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

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

| ORDER NUMBER | ORDER DATE | CUSTOMER NUMBER | SALES PERSON | PURCHASE ORDER NUMBER | SHIP VIA | SHIP DATE | TERMS |
|--------------|------------|-----------------|--------------|-----------------------|----------|-----------|--------|
| 100049766 | 11-11-08 | CUS001410 | 099325 | 010013704 | FXG | 12-03-08 | Net 30 |

ITEM DESCRIPTION

QUANTITY ORDERED QUANTITY SHIPPED STK UNIT PRICE UNIT DISCOUNT EXTENDED PRICE

Scale: 24 % | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

Feature Occurrences: 0 Total Classes: 0 Class Trained Items: 0 Total Weight: 0

Feature Count CWF CTC CF TF IDF Weight

OK Cancel

Pattern Editor Properties

Value Pattern

1 (@Number.{2,12})

Look Ahead Pattern (click to expand)

Look Behind Pattern (click to expand)

Output Format

{1:Number}

OK: ([0-9]{0}OoGgSsBbUuD){2,12})

Batch: Invoices

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

Image Text

Enid Parts

INVOICE

INVOICE DATE: 12-05-08 INVOICE #: CSI/10076138 PAGE: 1

PH: (405) 555-1212 FAX: (405) 444-2121

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

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

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

| ORDER NUMBER | ORDER DATE | CUSTOMER NUMBER | SALES PERSON | PURCHASE ORDER NUMBER | SHIP VIA | SHIP DATE | TERMS |
|--------------|------------|-----------------|--------------|-----------------------|----------|-----------|--------|
| 100049766 | 11-11-08 | CUS001410 | 099325 | 010013704 | FXG | 12-03-08 | Net 30 |

ITEM DESCRIPTION

QUANTITY ORDERED QUANTITY SHIPPED STK UNIT PRICE UNIT DISCOUNT EXTENDED PRICE

Scale: 34 % | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

Results (64) Confidence Page No Index Length

| | | | | |
|----|-------|---|----|---|
| 01 | 100 % | 1 | 14 | 2 |
| 01 | 100 % | 1 | 27 | 2 |
| 12 | 100 % | 1 | 40 | 2 |
| 05 | 100 % | 1 | 43 | 2 |

File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources)
 - _Global
 - Extractors
 - Lexicons
 - Ovemedes
 - [Enid] Sales Tax (data model)
 - Acme
 - Enid
 - Express
 - Spartan
 - Standard
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Field Class Weightings Contents Advanced

Save Cancel Run Extraction

Image View Text View

Enid Parts

INVOICE

INVOICE DATE: 12-05-08 INVOICE #: CSI/10076138 PAGE: 1

PH: (405) 555-1212 FAX: (405) 444-2121

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

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

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

| ORDER NUMBER | ORDER DATE | CUSTOMER NUMBER | SALES PERSON | PURCHASE ORDER NUMBER | SHIP VIA | SHIP DATE | TERMS |
|--------------|------------|-----------------|--------------|-----------------------|----------|-----------|--------|
| 100049766 | 11-11-08 | CUS001410 | 099325 | 010013704 | FXG | 12-03-08 | Net 30 |

ITEM DESCRIPTION

QUANTITY ORDERED QUANTITY SHIPPED STK UNIT PRICE UNIT DISCOUNT EXTENDED PRICE

Scale: 24 % | 2548px x 3287px | 8.49" x 10.96" | 300 DPI | Black & White

Feature Occurrences: 0 Total Classes: 0 Class Trained Items: 0 Total Weight: 0

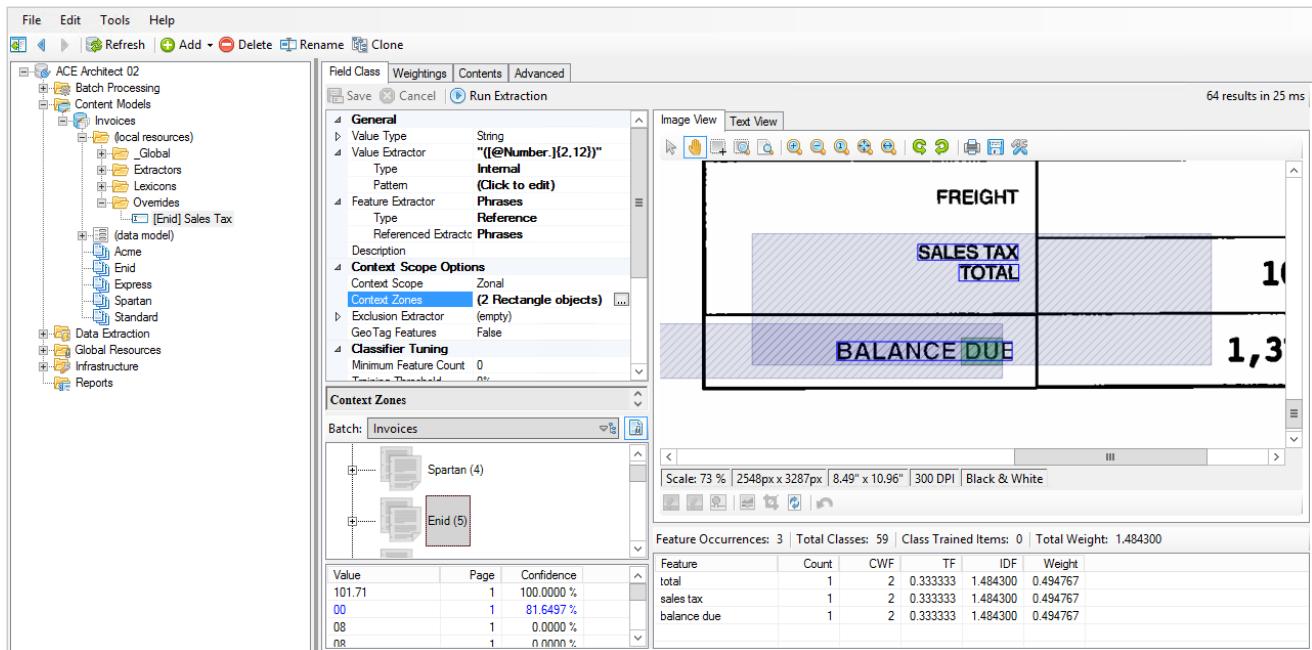
Feature Count CWF CTC CF TF IDF Weight

> Step

1. Click to edit the Context Zones property
2. In the Context Zones window, delete one of the zones
3. Edit the dimensions of the remaining zone:

| | Zone |
|--------|------|
| Left | -2.4 |
| Top | -0.2 |
| Right | 0.25 |
| Bottom | 0.1 |

4. Click OK to close the Context Zones window
5. In the Output section, set the Minimum Confidence property to 90%



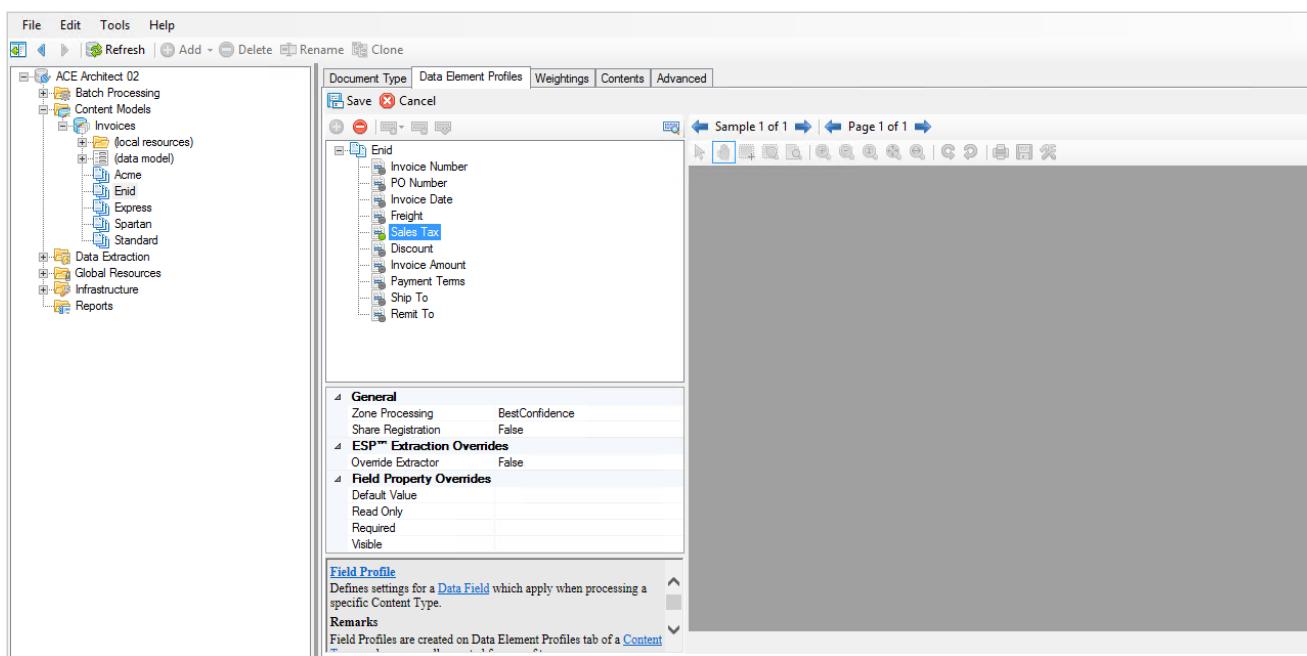
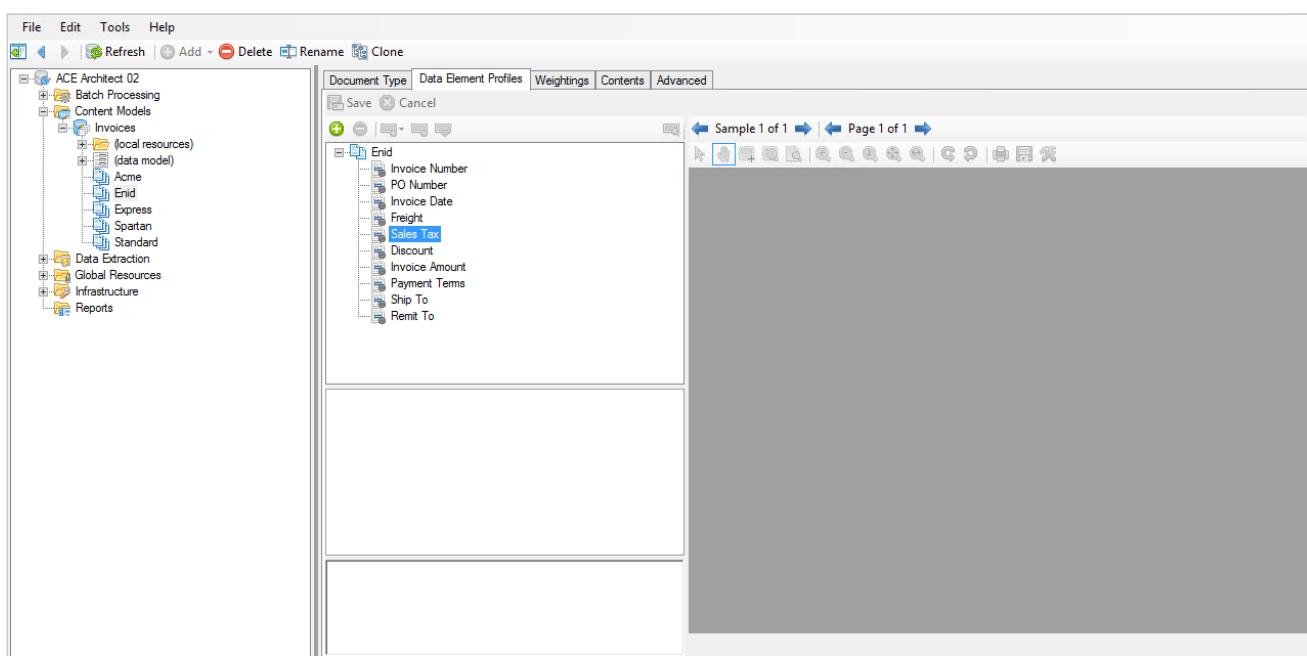
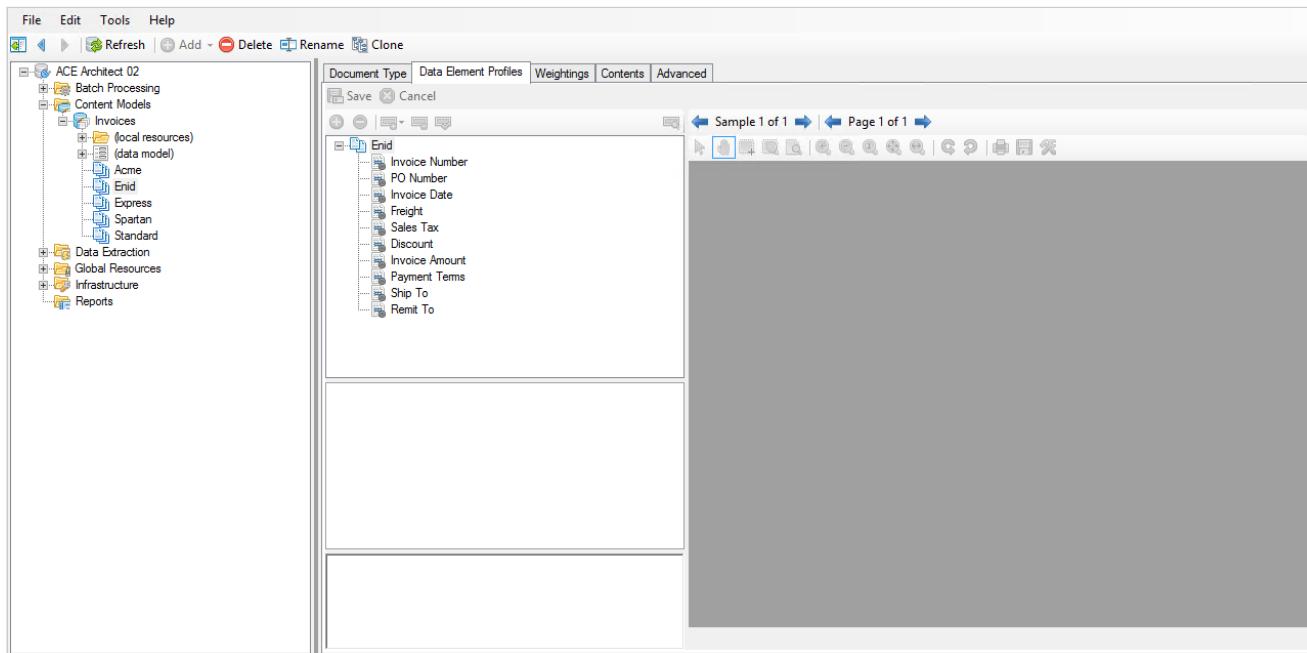
The screenshot shows the ACE Architect 02 interface. On the left, the navigation tree includes 'ACE Architect 02', 'Batch Processing', 'Content Models', 'Invoices' (selected), 'Extractors', 'Lexicons', 'Overrides', and 'Reports'. In the center, the 'Field Class' tab is active, showing a 'Value Type' of 'String' and a 'Value Extractor' pattern of '"(#[@Number.](2,12))"'. Below this is the 'Context Zones' dialog, which defines a zone for 'Zone 2' with coordinates '(-2.4, -0.2):(0.25, 0.1)'. To the right is the 'Image View' showing a portion of an invoice with the value '101.71' highlighted. At the bottom, a feature occurrence table lists 'total', 'sales tax', and 'balance due'.

This screenshot shows the ACE Architect 02 interface with a different configuration. The 'Context Scope Options' are set to 'Zonal' with a single 'Context Zone' of '(1 Rectangle)'. The 'Classifier Tuning' section has a 'Minimum Confidence' of 90%. The 'Output' section shows a 'Minimum Confidence' of 90%, 'Collation Method' of 'Individual', and 'Order By' of 'Confidence'. The 'Feature Occurrences' table at the bottom shows 'total' and 'sales tax' features.

Step

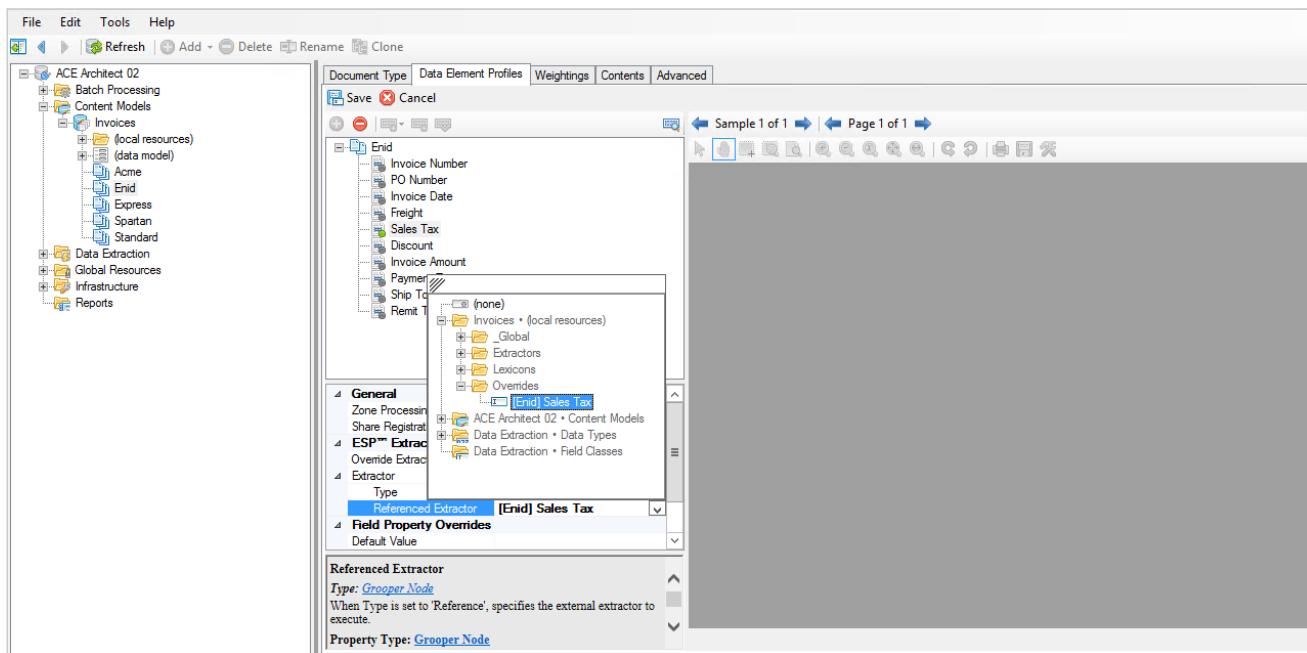
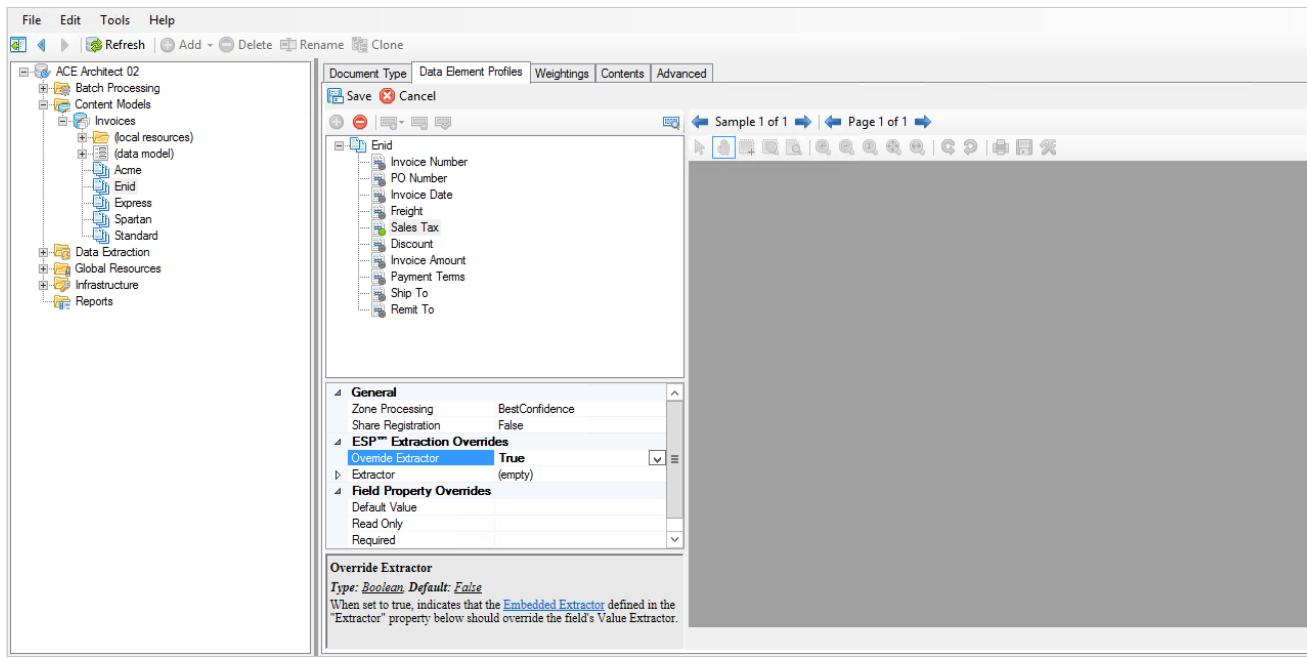
1. Navigate to the (root) > Content Models > Invoices > Enid Document Type
2. Click on the Data Element Profiles tab
3. Click on the Sales Tax Data Field and click the + button in the toolbar

This screenshot shows the 'Data Element Profiles' tab for the 'Enid' document type. The 'General' section includes 'Allow Training' (True), 'Page Scope - Data Extraction' (Unlimited), 'Base Content Type' (Document), and 'Description'. The 'Classification Rules' section contains a positive extractor rule for 'enid parts'. The 'Separation' section includes 'Pagination' (Structured) and 'Combine Contiguous Documents' (False). On the right, the 'Data Model Preview' shows fields for 'Invoice Number', 'PO Number', 'Invoice Date', 'Freight', 'Sales Tax', 'Discount', 'Invoice Amount', 'Payment Terms', 'Ship To', and 'Remit To'. A note at the bottom explains that document types represent distinct types of documents and are assigned to batches for classification.



Step

1. In the properties panel, set the ESP™ Extraction Overrides > Override Extractor property to **True**
2. Expand Extractor and set the Extractor > Type property to Reference
3. Set the Referenced Extractor property to the Invoices • (local resources) > Overrides > **[Enid] Sales Tax** Field Class



Data Model Adjustments

Step

1. Navigate to the Ship To Data Field
2. In the Appearance section, set the Display Width property to **110**
3. In the Behavior section, set the Multi Line property to **Enabled**
4. Expand Multi Line and set the Multi Line Height property to **45**
5. Set the Word Wrap property to **True**

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources) (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **Cancel** **Test Extraction**

General

Appearance

Alignment: Left
Display Width: 110
Error Color: 255, 192, 192
Background Color: Window
Foreground Color:WindowText
Visible: True
Label Position: Left
Label Alignment: MiddleLeft

Behavior

FieldSense™: Disabled
Input Mask: Disabled
Multi Line: Enabled
Tooltip: (Text)
Read Only: False
Required: True
Sticky: False
Requires Validation: False
Character Cascading: Normal

Display Width

Batch: Invoices

Invoices

- Acme (1)
- Standard (2)

Ship To

Data Field Preview

Document View **Text View** **Instance View**

ACME | INTERNATIONAL

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

Invoice

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

Page 1 / 01

P.O. Reference

| | | | |
|-------------------|-------------|------------------|---------------------|
| P.O. number: | 201017081 | Date: | 11/14/2008 |
| Customer number: | 18003405 | Invoice number: | 74451405 |
| Customer contact: | 18102943 | Ship-to Number: | 33241729 |
| Currency: | USD | Sales order no.: | 5515010 |
| Invoice amount: | 1,227.13 | Delivery date: | 05/15/09 |
| Payment terms: | Net 30 days | Freight carrier: | |
| | | Tracking number: | 1234567890301034108 |

Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources) (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **Cancel** **Test Extraction**

General

Appearance

Alignment: Left
Display Width: 110
Error Color: 255, 192, 192
Background Color: Window
Foreground Color:WindowText
Visible: True
Label Position: Left
Label Alignment: MiddleLeft

Behavior

FieldSense™: Disabled
Input Mask: Disabled
Multi Line: Enabled
Tooltip: (Text)
Read Only: False
Required: True
Sticky: False
Requires Validation: False
Character Cascading: Normal

Multi Line

Batch: Invoices

Invoices

- Acme (1)
- Standard (2)

Ship To

Data Field Preview

Document View **Text View** **Instance View**

ACME | INTERNATIONAL

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

Invoice

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

Page 1 / 01

P.O. Reference

| | | | |
|-------------------|-------------|------------------|---------------------|
| P.O. number: | 201017081 | Date: | 11/14/2008 |
| Customer number: | 18003405 | Invoice number: | 74451405 |
| Customer contact: | 18102943 | Ship-to Number: | 33241729 |
| Currency: | USD | Sales order no.: | 5515010 |
| Invoice amount: | 1,227.13 | Delivery date: | 05/15/09 |
| Payment terms: | Net 30 days | Freight carrier: | |
| | | Tracking number: | 1234567890301034108 |

Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

ACE Architect 02

- Batch Processing
- Content Models
 - Invoices
 - (local resources) (data model)
 - Invoice Number
 - PO Number
 - Invoice Date
 - Freight
 - Sales Tax
 - Discount
 - Invoice Amount
 - Payment Terms
 - Ship To
 - Remit To
- Data Extraction
- Global Resources
- Infrastructure
- Reports

Data Field **Contents** **Advanced**

Save **Cancel** **Test Extraction**

General

Appearance

Alignment: Left
Display Width: 110
Error Color: 255, 192, 192
Background Color: Window
Foreground Color:WindowText
Visible: True
Label Position: Left
Label Alignment: MiddleLeft

Behavior

FieldSense™: Disabled
Input Mask: Disabled
Multi Line: Enabled
Multi Line Height: 45
Word Wrap: True
Tooltip: (Text)
Read Only: False
Required: True
Sticky: False

Rubberband OCR Profile

Batch: Invoices

Invoices

- Acme (1)
- Standard (2)

Ship To

Data Field Preview

Document View **Text View** **Instance View**

ACME | INTERNATIONAL

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

Invoice

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

Page 1 / 01

P.O. Reference

| | | | |
|-------------------|-------------|------------------|---------------------|
| P.O. number: | 201017081 | Date: | 11/14/2008 |
| Customer number: | 18003405 | Invoice number: | 74451405 |
| Customer contact: | 18102943 | Ship-to Number: | 33241729 |
| Currency: | USD | Sales order no.: | 5515010 |
| Invoice amount: | 1,227.13 | Delivery date: | 05/15/09 |
| Payment terms: | Net 30 days | Freight carrier: | |
| | | Tracking number: | 1234567890301034108 |

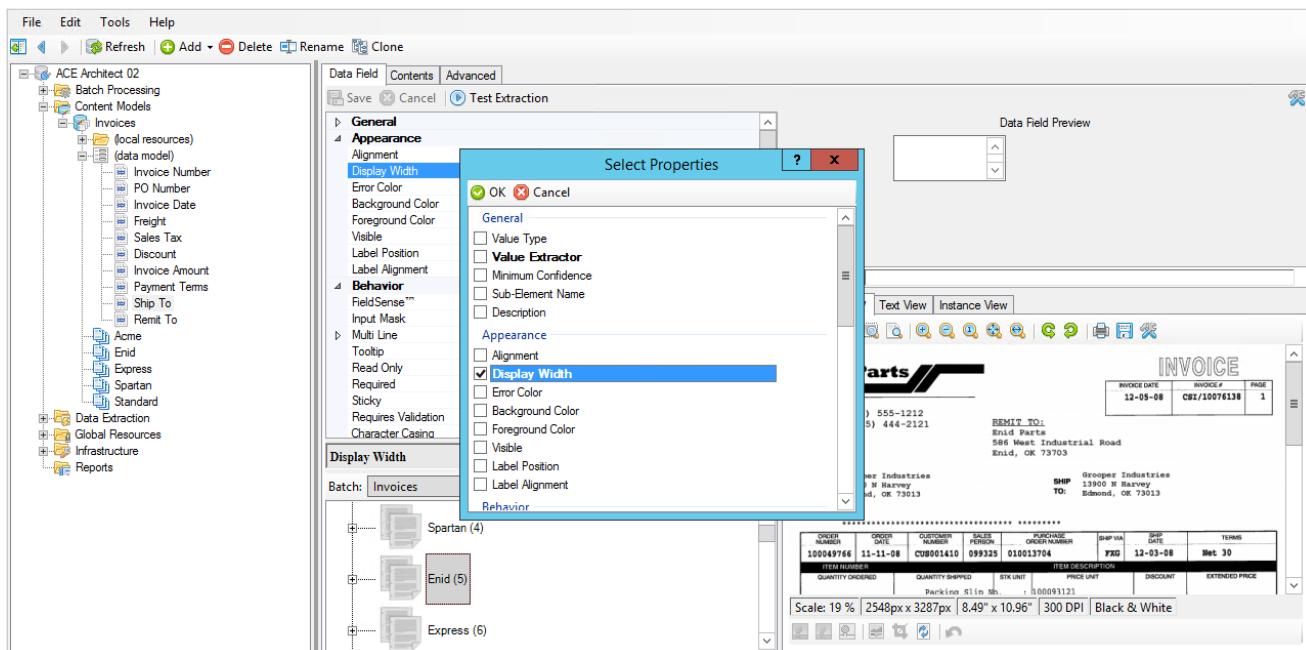
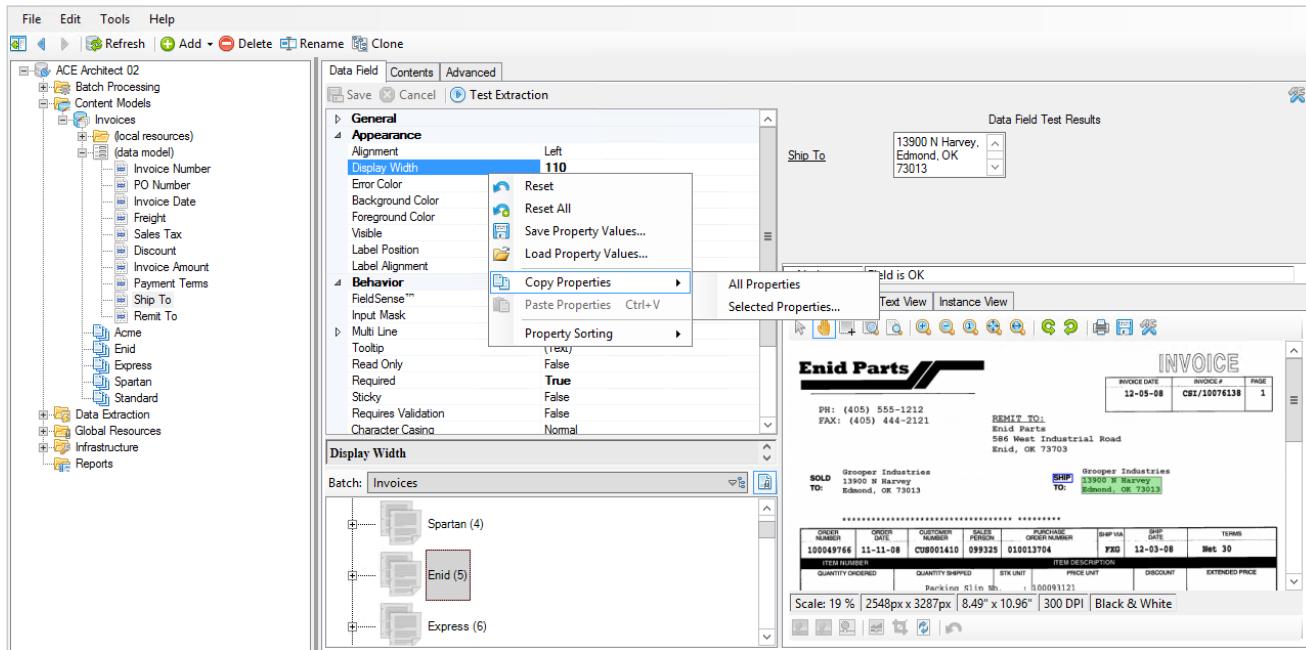
Scale: 19% | 2545px x 3285px | 8.48" x 10.95" | 300 DPI | Black & White

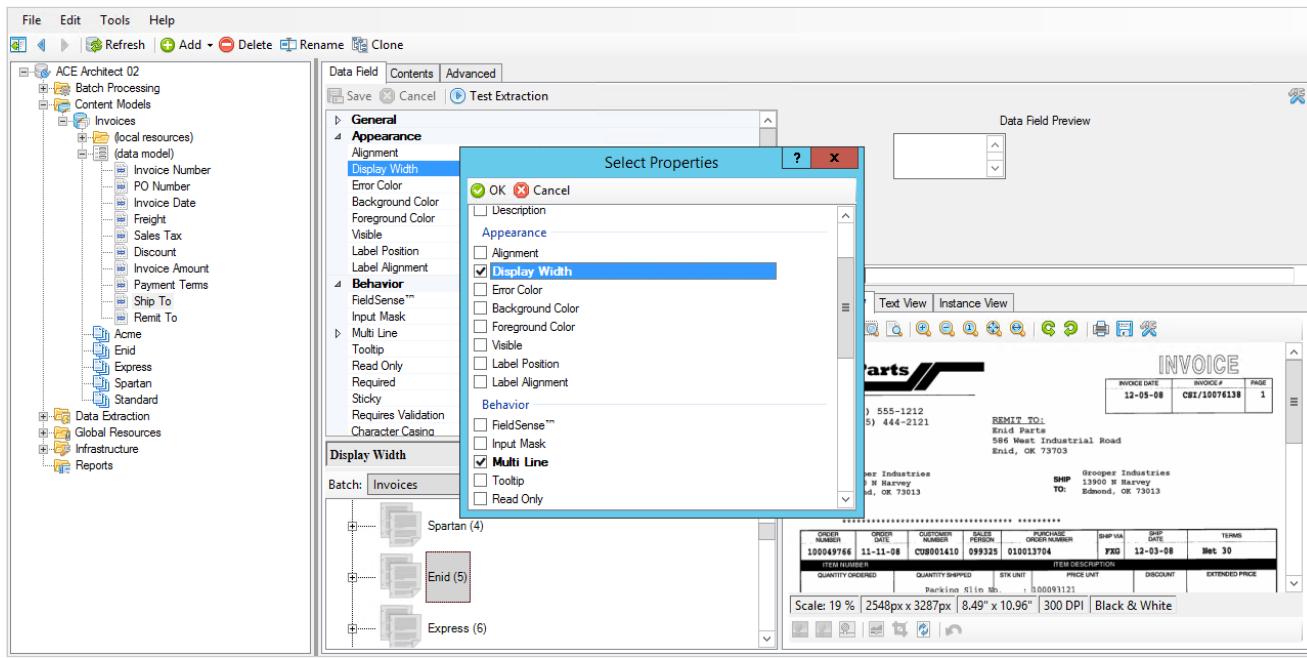
Step

1. Right click on any of the property names and select Copy Properties > Selected Properties...
2. In the Select Properties window, check the boxes next to the properties to copy:

Display Width

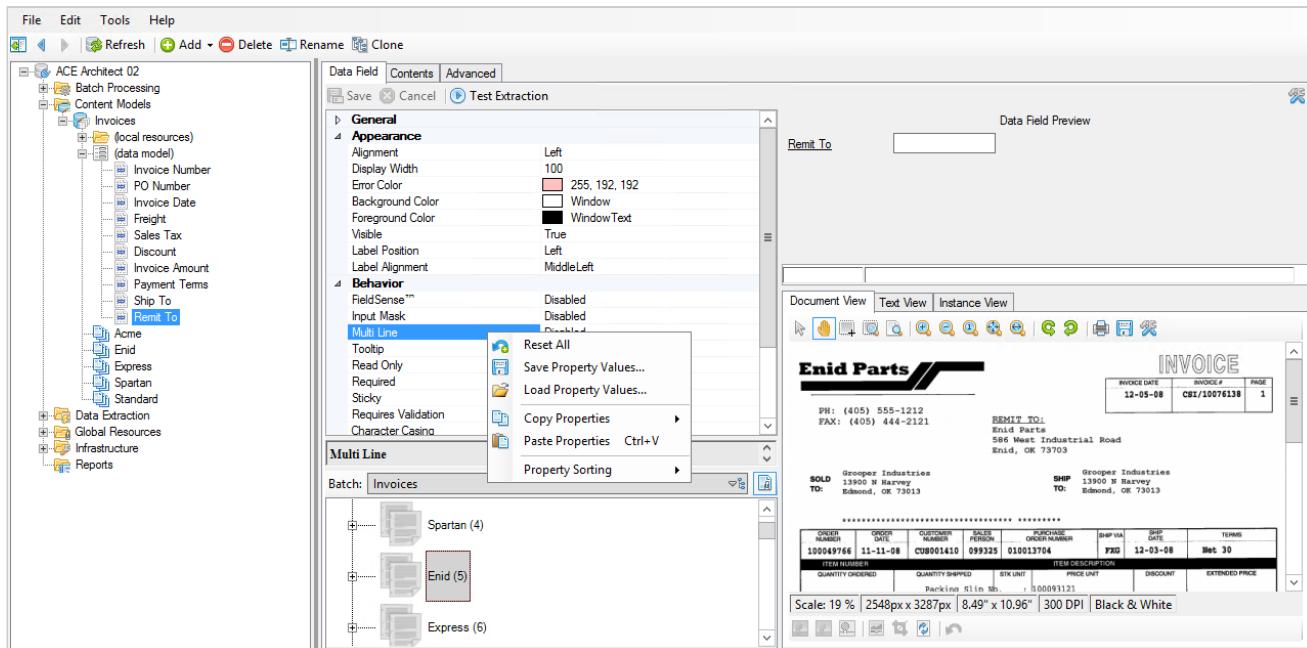
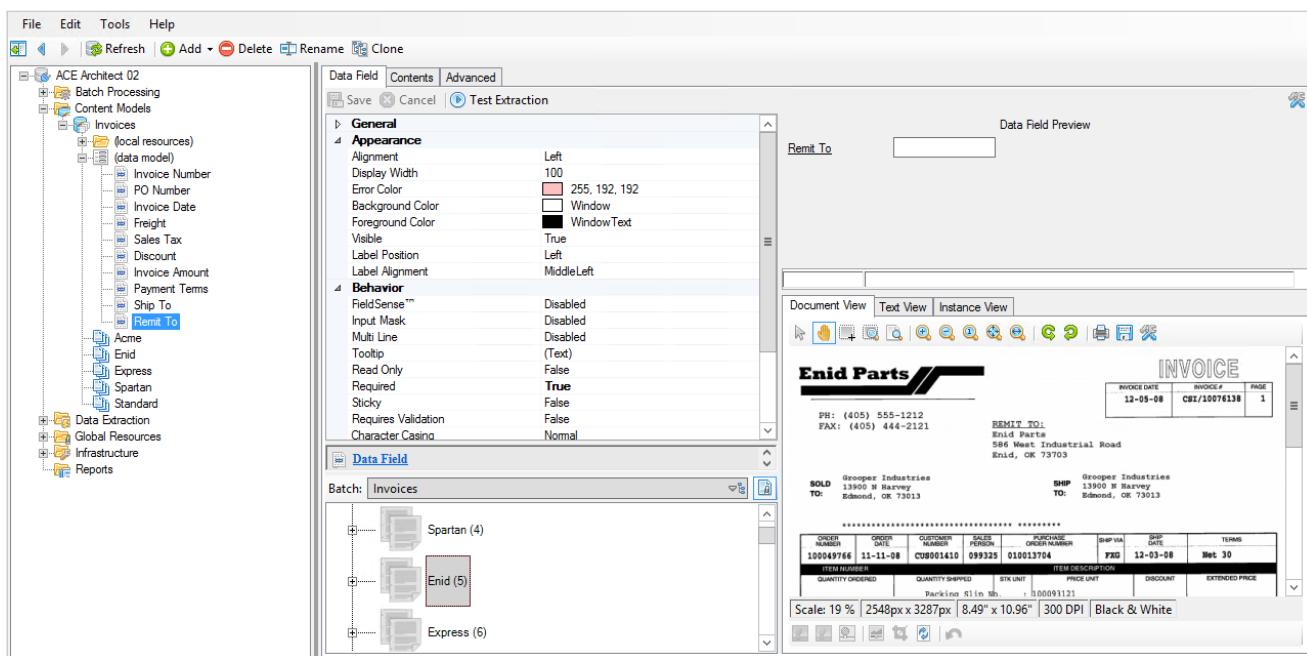
Multi Line

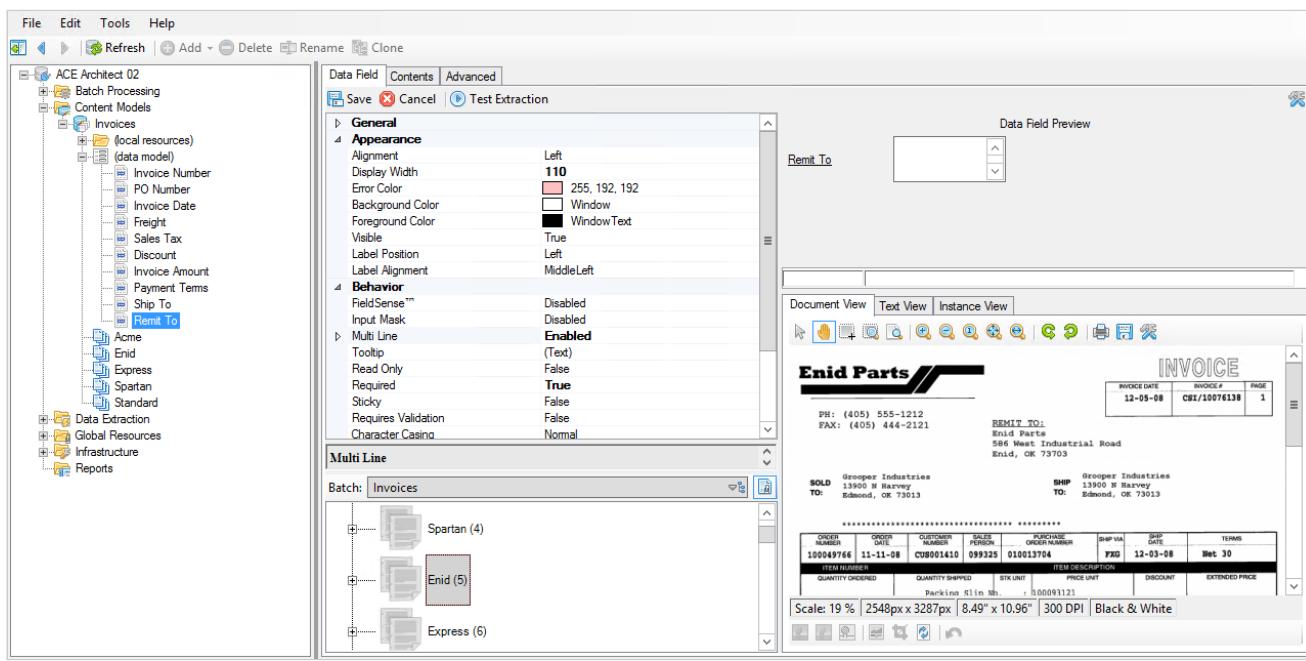




Step

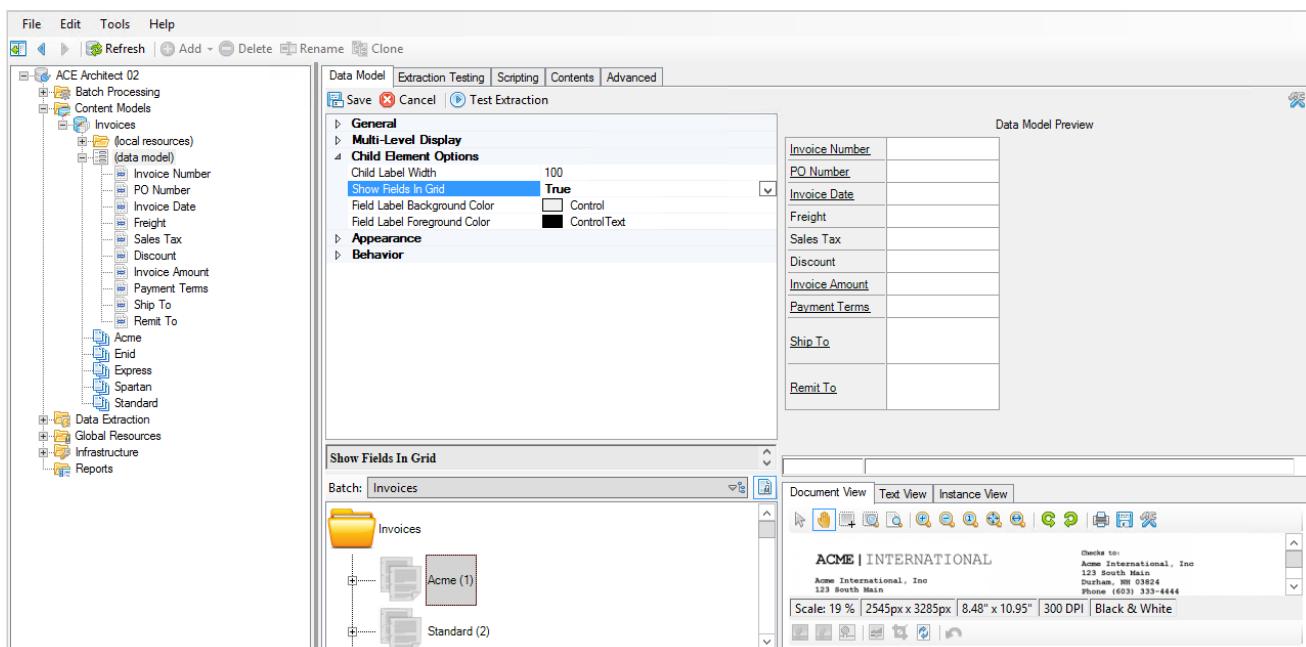
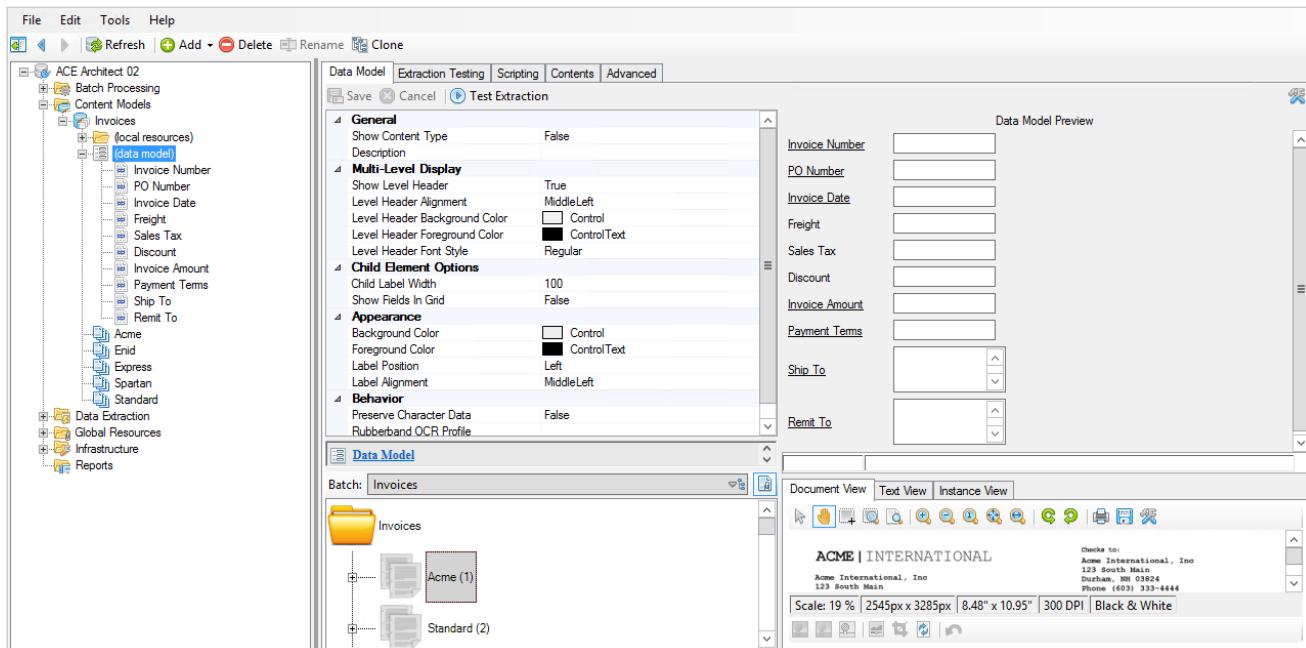
1. Navigate to the Remit To Data Field
2. Right click on any of the property names and select Paste Properties





Step

1. Navigate to the Data Model itself
2. In the Child Element Options section, set the Show Fields In Grid property to **True**
3. In the toolbar, click the Test Extraction button to test all fields against a document



File Edit Tools Help

Refresh Add Delete Rename Clone

ACE Architect 02

Batch Processing Content Models Invoices (local resources) (data model)

- Invoice Number
- PO Number
- Invoice Date
- Freight
- Sales Tax
- Discount
- Invoice Amount
- Payment Terms
- Ship To
- Remit To

Acme End Express Spartan Standard

Data Extraction Global Resources Infrastructure Reports

Data Model Extraction Testing Scripting Contents Advanced

Save Cancel Test Extraction

General Multi-Level Display Child Element Options

Child Label Width 100 Show Fields In Grid True

Field Label Background Color Control Field Label Foreground Color ControlText

Appearance Behavior

Invoice Number 74451405
PO Number 201017081
Invoice Date 11/14/2008
Freight \$4.20
Sales Tax \$0.00
Discount \$246.13
Invoice Amount \$1,727.13
Payment Terms Net 30
Ship To 13900 N Harvey, Edmond, OK 73013
Remit To 123 South Main, Durham, NH 03824

Show Fields In Grid

Batch: Invoices

Invoices Acme (1)
Standard (2)

Data Model Test Results

No Issues Field is OK

Document View Text View Instance View

13900 N Harvey Edmond, OK 73013 405-507-7009 13900 N Harvey Edmond, OK 73013 405-507-7009

Your Reference Our Reference

Scale: 19% 254px x 328px 8.48" x 10.95" 300 DPI Black & White

The screenshot shows the ACE Architect software interface. On the left, there's a navigation tree with sections like 'ACE Architect 02', 'Batch Processing', 'Content Models', 'Invoices' (selected), 'Data Extraction', 'Global Resources', 'Infrastructure', and 'Reports'. The 'Invoices' section is expanded, showing fields such as 'Invoice Number', 'PO Number', 'Invoice Date', 'Freight', 'Sales Tax', 'Discount', 'Invoice Amount', 'Payment Terms', 'Ship To', and 'Remit To'. Below this are icons for 'Acme', 'End', 'Express', 'Spartan', and 'Standard'. The top menu bar includes 'File', 'Edit', 'Tools', 'Help', and various icons for refresh, add, delete, rename, and clone. The main workspace has tabs for 'Data Model', 'Extraction Testing', 'Scripting', 'Contents', and 'Advanced'. The 'Extraction Testing' tab is active, showing a 'Test Extraction' dialog with sections for 'General', 'Multi-Level Display', 'Child Element Options' (with 'Show Fields In Grid' set to 'True'), 'Appearance', and 'Behavior'. A preview pane on the right shows 'Data Model Test Results' with a grid of invoice data and a preview of the document with 'Text View', 'Instance View', and 'Document View' tabs. The preview pane also shows document details like '13900 N Harvey Edmond, OK 73013 405-507-7009' and printing options like 'Scale: 19%', '254px x 328px', '8.48" x 10.95"', '300 DPI', and 'Black & White'.