



IBM Integration Bus

Message Modeling with DFDL

Lab 4

Record-oriented, tagged, delimited text (advanced)

June, 2013

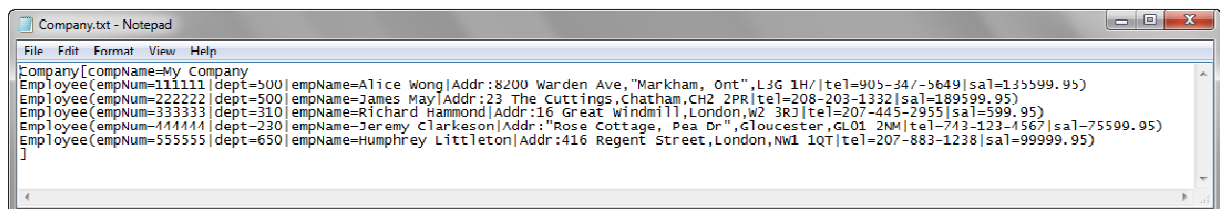
Hands-on lab built at product
code level Version 9.0

1. Introduction

1.1 Lab scenario

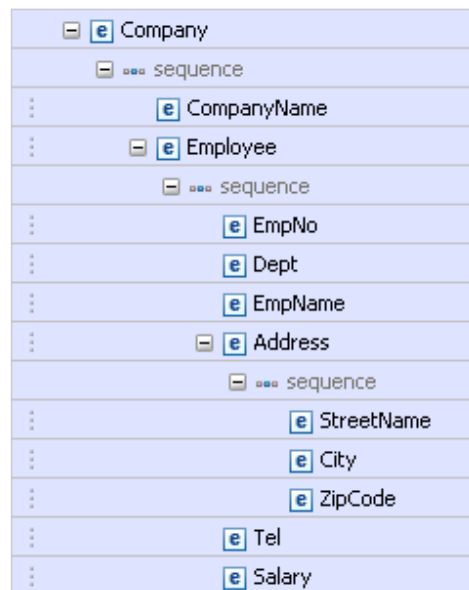
In this lab you are going to extend the message model you created in the previous (basic) lab, Lab3.

In the basic lab, you created a message model to parse this data file:



```
Company[compName=My Company
Employee(empNum=111111|dept=500|empName=Alice Wong|Addr:8200 Warden Ave,"Markham, Ont",L3G 1H|tel=905-347-5649|sal=135599.95)
Employee(empNum=222222|dept=500|empName=James May|Addr:23 The Cuttings,Chatham,CH2 2PR|tel=208-203-1332|sal=189599.95)
Employee(empNum=333333|dept=310|empName=Richard Hammond|Addr:16 Great Windmill,London,W2 3RJ|tel=207-445-2955|sal=599.95)
Employee(empNum=444444|dept=230|empName=Jeremy Clarkson|Addr:"Rose Cottage, Pea Dr",Gloucester,GL01 2NM|tel=743-123-4567|sal=75599.95)
Employee(empNum=555555|dept=650|empName=Humphrey Littleton|Addr:416 Regent Street,London,NW1 1QT|tel=207-883-1238|sal=99999.95)
]
```

So you defined the following DFDL structure like this:

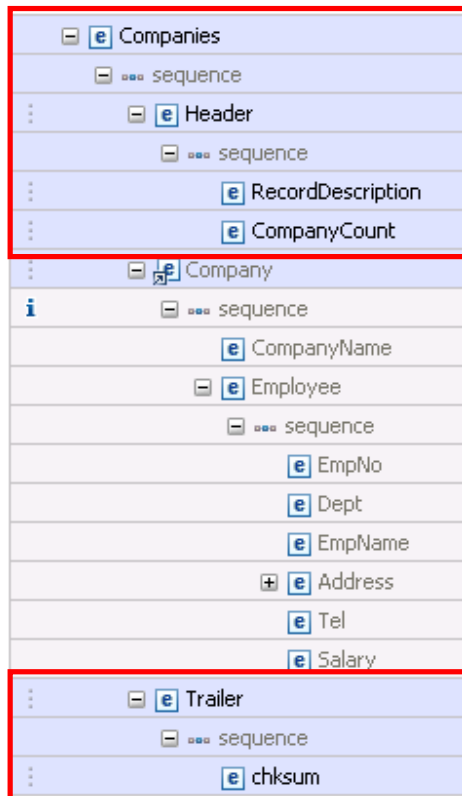


In this lab, you will extend this DFDL structure so that it can parse the following data file which contains multiple Company records:

```

Companies.txt - Notepad
File Edit Format View Help
Header{recDesc:My Company records,compcount:5}
Company[compName=BBK
Employee(empNum=111111|dept=500|empName=Alice Wong|Addr:8200 Warden Ave,"Markham, Ont",L3G 1H7|tel=905-347-5649|sal=135599.95)
Employee(empNum=222222|dept=500|empName=James May|Addr:23 The Cuttings,Chatham, CH2 2PR|tel=208 203 1332|sal=6189599.95)
Employee(empNum=333333|dept=310|empName=Richard Hammond|Addr:16 Great Windmill St,London,W2 3RJ|tel=207-445-2955|sal=599.95)
Employee(empNum=444444|dept=230|empName=Jeremy Clarkson|Addr:"Rose Cottage, Pea Ur",Gloucester,GL01 2NM|tel=743-123-4567|sal=5599.95)
Employee(empNum=555555|dept=650|empName=Humphrey Littleton|Addr:416 Regent Street,London,Nw1 1QT|tel=207-883-1238|sal=99999.95)
]
Company[compName=IBM
Employee(empNum=111111|dept=9876|empName=Arnold Buzby|Addr:1000 The Close,Winchester,L3G 1H7|tel=905-345-5649|sal=23.54)
Employee(empNum=222222|dept=2350|empName=Digby Jones|Addr:1 Porstmouth Rd,Southampton,CH2 2PR|tel=208-203-1332|sal=599.95)
]
Company[compName=Big Bank
Employee(empNum=111111|dept=1|empName=Mr Big|Addr:99 Sicillian Dr,"Palermo, NY",L3G 1H7|tel=905-347-5649|sal=4599.00)
Employee(empNum=000001|dept=1|empName=Homer Simpson|Addr:Two's Complement,Springfield,1011001|tel=208-203-1332|sal=189599.94)
Employee(empNum=333333|dept=2|empName=Lucy Wetherell|Addr:23 Gas St,Bolton,W2 3RJ|tel=207-445-2955|sal=599.95)
]
Company[compName=Huge Store
Employee(empNum=111111|dept=20|empName=George Formby|Addr:1 HotPot Rd,Lancashire,L3G 1H7|tel=905-347-5649|sal=85599.95)
Employee(empNum=222222|dept=18|empName=Ivor Engine|Addr:1234 London Rd,LtAngollen,CH2 2PR|tel=208-203-1332|sal=9.95)
]
Company[compName=Corner Store
Employee(empNum=100001|dept=4456|empName=Captain Black|Addr:The Fire Station,Trumpton,L3G 1H7|tel=905 347 5649|sal=12345.95)
Employee(empNum=100002|dept=4429|empName=Captain Pugwash|Addr:The Black Pig,Smugglers Cove,CH2 2PR|tel=208-203-1332|sal=654321.91)
Employee(empNum=100003|dept=4420|empName=Lady Penelope|Addr:Creighton-Ward Mansion,Buckinghamshire,W2 3RJ|tel=207-445-2955|sal=599.95)
Employee(empNum=100004|dept=4483|empName=Jeff Tracy|Addr:Tracy Island,Pacific Ocean,GL012NM|tel=743-123-4567|sal=75599.23)
Employee(empNum=100005|dept=4400|empName=Gordon Tracy|Addr:Tracy Island,Pacific Ocean,Nw1 1QT|tel=207-883-1238|sal=666.67)
]
Trailer{chksum:1234567890}
    
```

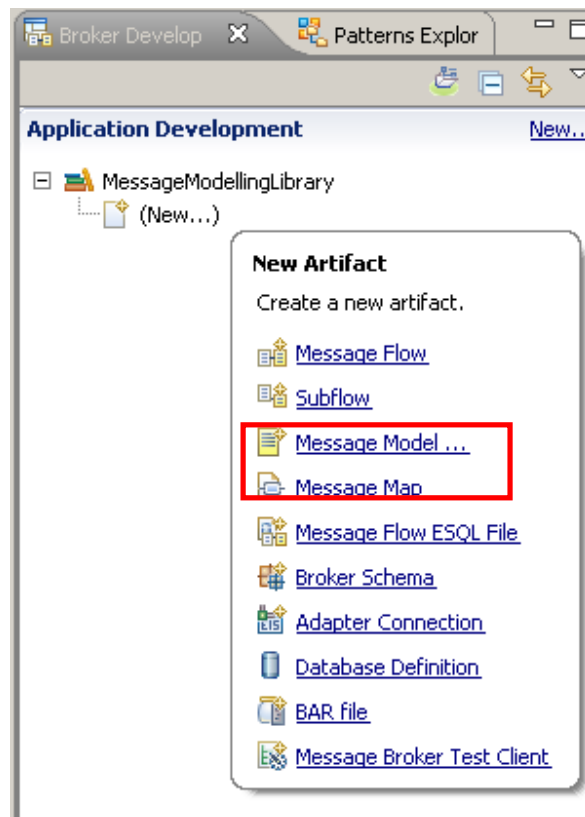
So you will need to create a structure like this:



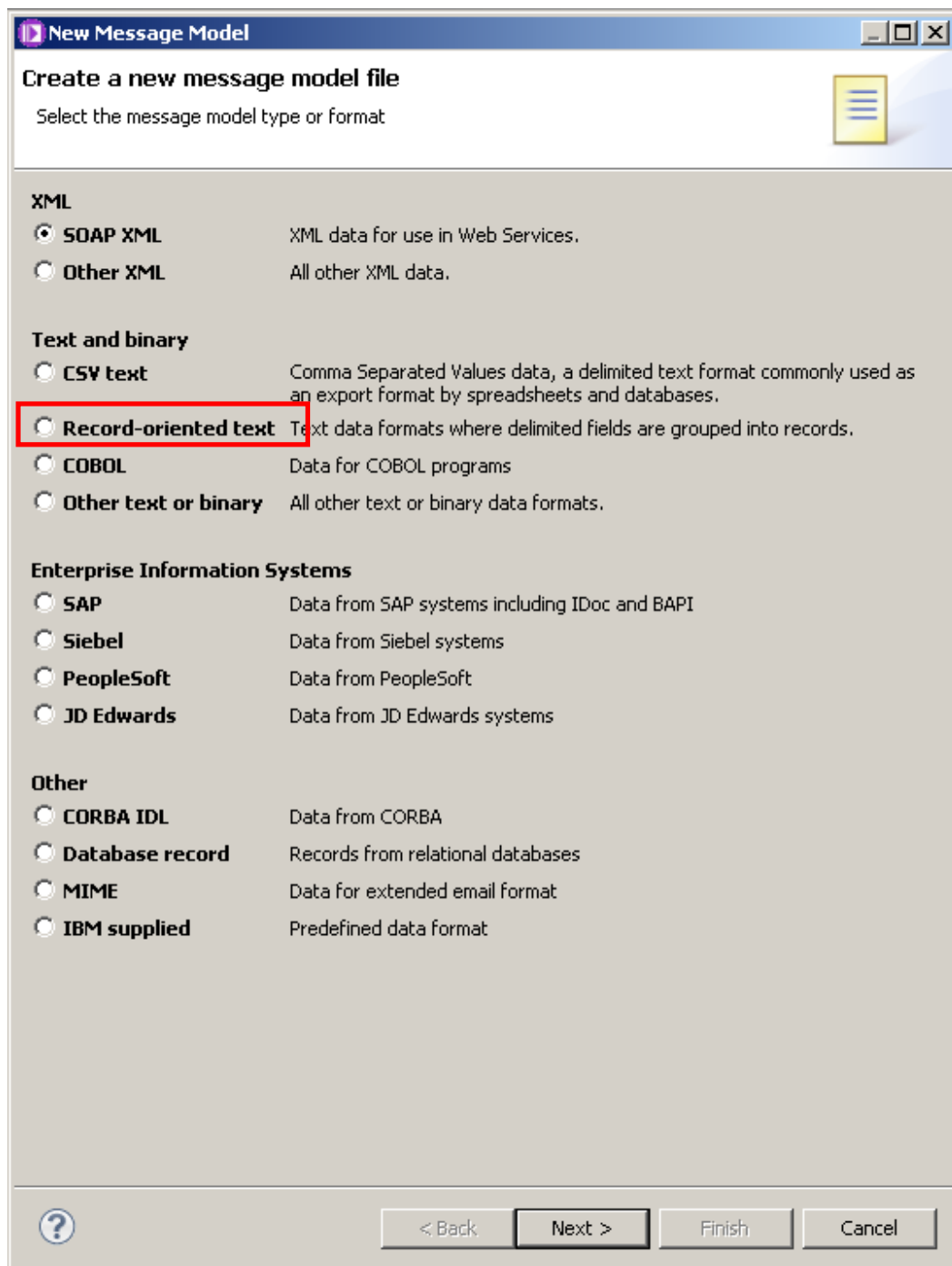
Then you are going to test it against a corresponding data file.

2. Create the Message Model in the Library

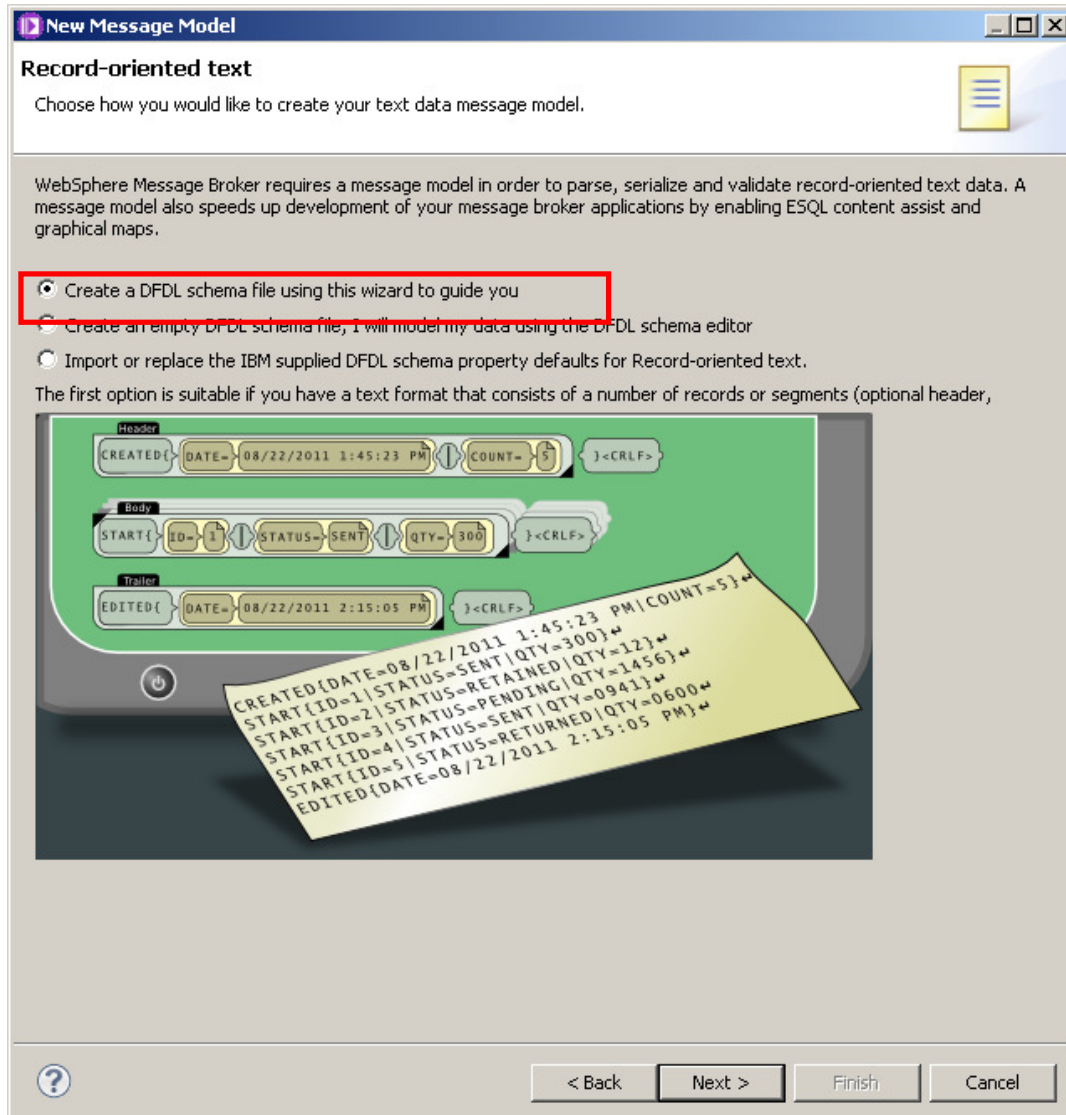
1. In the "MessageModellingLibrary library, click New->Message Model to create a new DFDL Schema.



2. In the "New Message Model" wizard, select "Record-oriented text" and click Next.



- From the wizard, select "Create a DFDL schema file using the wizard to guide you" and click Next.



4. Enter "Companies" as the DFDL Schema file name and click Next.

New Message Model

Create a Data Format Description Language (DFDL) Schema

Specify the location and name of the DFDL schema, and specify the name of the message.

Project: Browse... New...

Folder: Browse... New...

DFDL schema file name:

Message name:

? < Back Next > Finish Cancel

Note that the Message name will auto complete based on the DFDL schema file name.

5. Leave the "End of record character" default value (carriage return, line feed).

Also leave the "The first record is a header" and "The last record is a trailer" checked.

In the Header fields tab, enter "Header{" as the Header initiator, and 2 as the "Number of fields"

New Message Model

Configure schema for data formatted as records and fields

Provide setting for new DFDL schema that represent record-oriented data.

Record settings

End of record character: Carriage Return & Line Feed - %CR;%LF;

(Blank records will be skipped)

The first record is a header

The last record is a trailer

Header fields | body fields | Trailer fields

Header initiator: Header{

Number of fields: 2

Field settings

Separated by: | - %#124; (UTF-8: 0x7C) (UTF-16: 0x007C)

Fixed length

All fields have an initiator

Create default values for fields

Encoding code page options:

Dynamic (provided to the processor by the application at runtime)

Fixed UTF-8

Global settings

Escape scheme: Default escape scheme

< Back Next > Finish Cancel

- Click on the "Trailer fields" tab, and enter "Trailer{" as the Trailer initiator and 1 as the Number of fields.

Click on the Finish button.

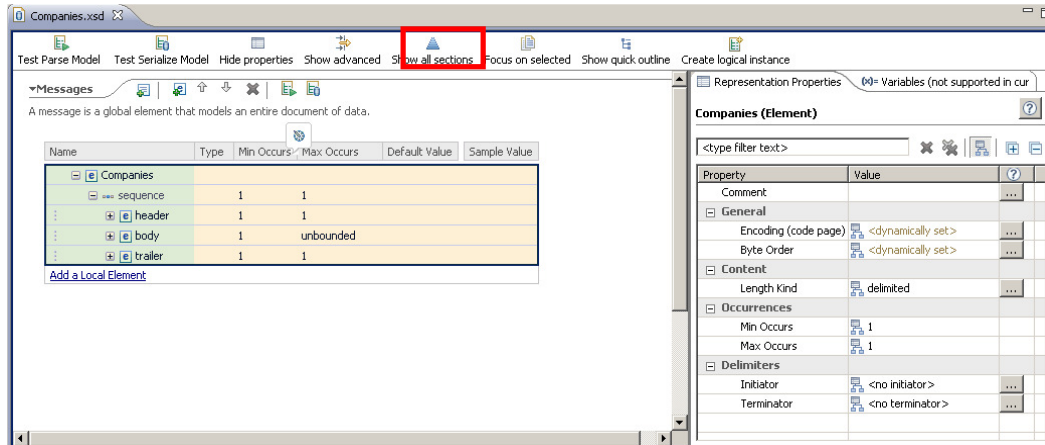
The screenshot shows the 'New Message Model' dialog box with the title 'Configure schema for data formatted as records and fields'. The 'Record settings' section includes a dropdown for 'End of record character' set to 'Carriage Return & Line Feed - %CR;%LF;', a note '(Blank records will be skipped)', and two checked checkboxes: 'The first record is a header' and 'The last record is a trailer'. The 'Trailer fields' tab is selected, and the 'Trailer initiator' is set to 'Trailer{' and 'Number of fields' is set to '1'. The 'Field settings' section has 'Separated by' set to '| - %#124; (UTF-8: 0x7C) (UTF-16: 0x007C)', 'Fixed length' is unselected, 'All fields have an initiator' is checked, and 'Create default values for fields' is unselected. The 'Encoding code page options' section has 'Dynamic' selected with the note '(provided to the processor by the application at runtime)' and 'Fixed' set to 'UTF-8'. The 'Global settings' section has 'Escape scheme' set to 'Default escape scheme'. At the bottom, there are buttons for '< Back', 'Next >', 'Finish', and 'Cancel'. The 'Finish' button is highlighted.

3.

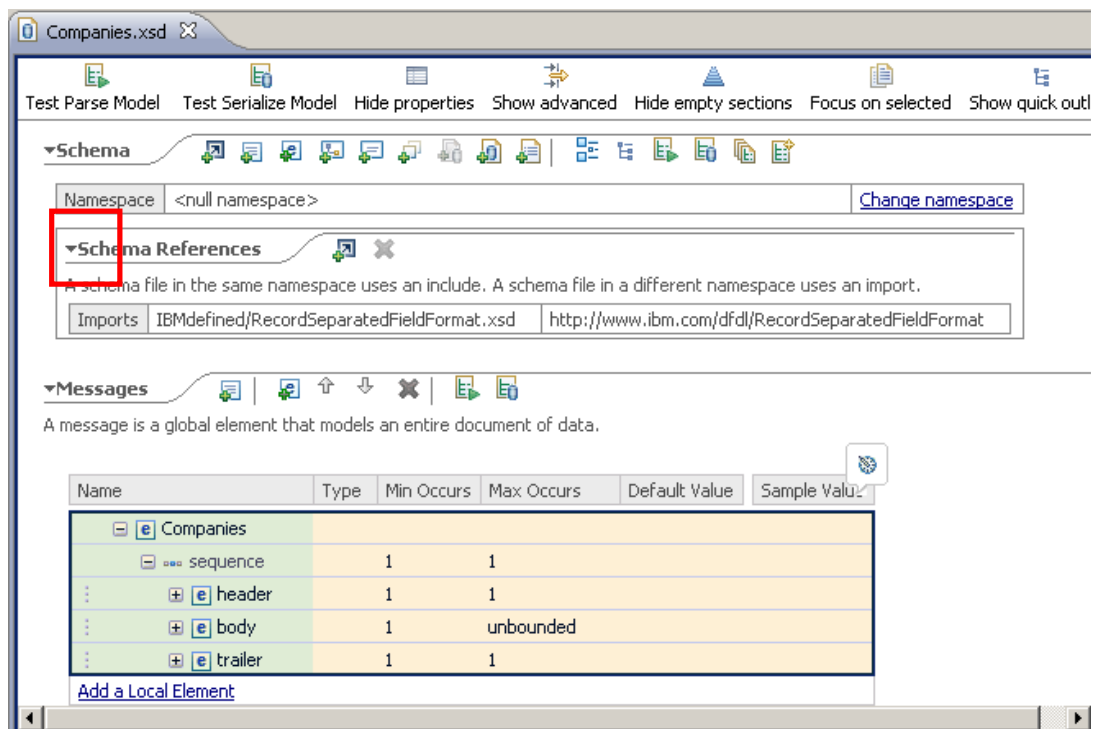
4. Refine the Message Model

1. The DFDL editor will open with the generated DFDL Schema.

Click on the "Show all sections" icon.

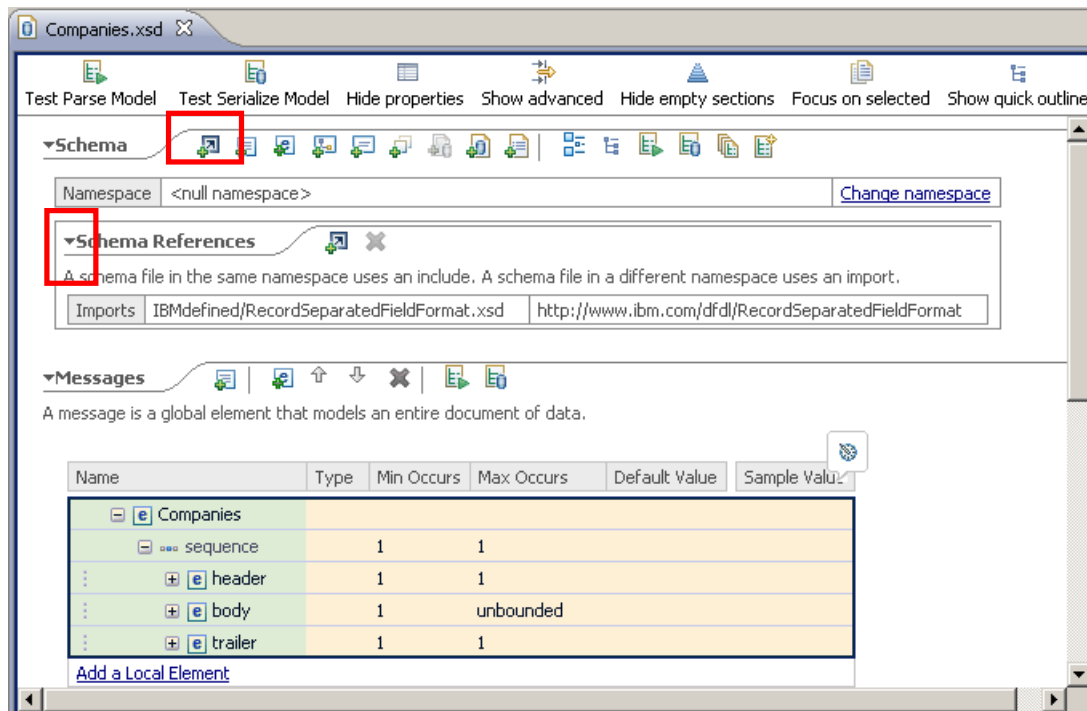


2. Click on the twisty next to the "Schema References" section to expand the references (Includes and Imports) of the DFDL Schema file.

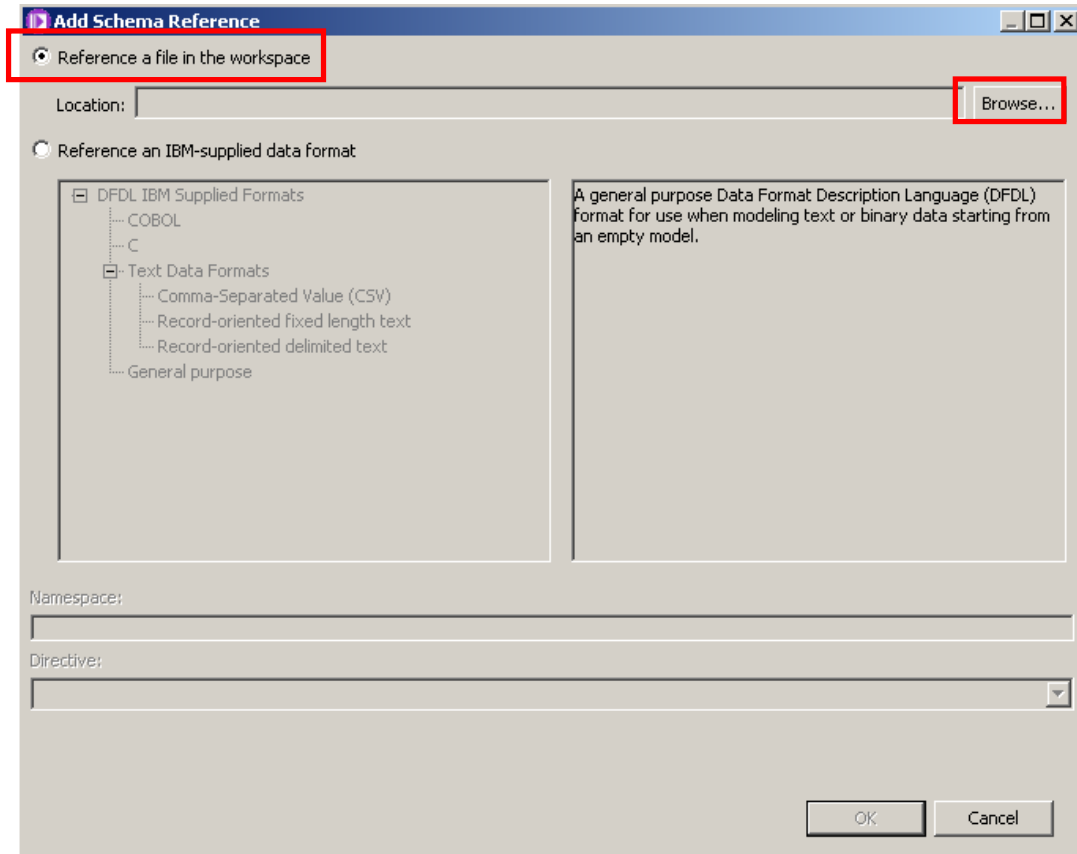


- This message model is going to build on the message model from the basic lab, by creating a reference to that model.

Click on the "Add a reference to another schema" icon.

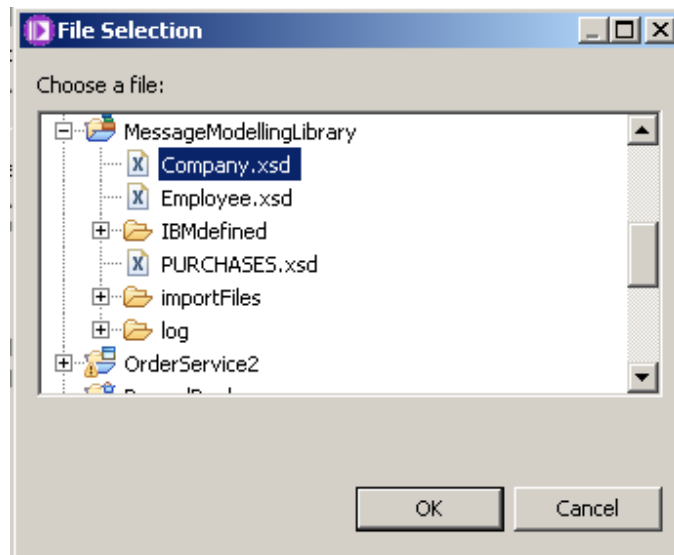


4. In the "Add Schema Reference" window, leave the default "Reference a file in the workspace" option and click the Browse button.

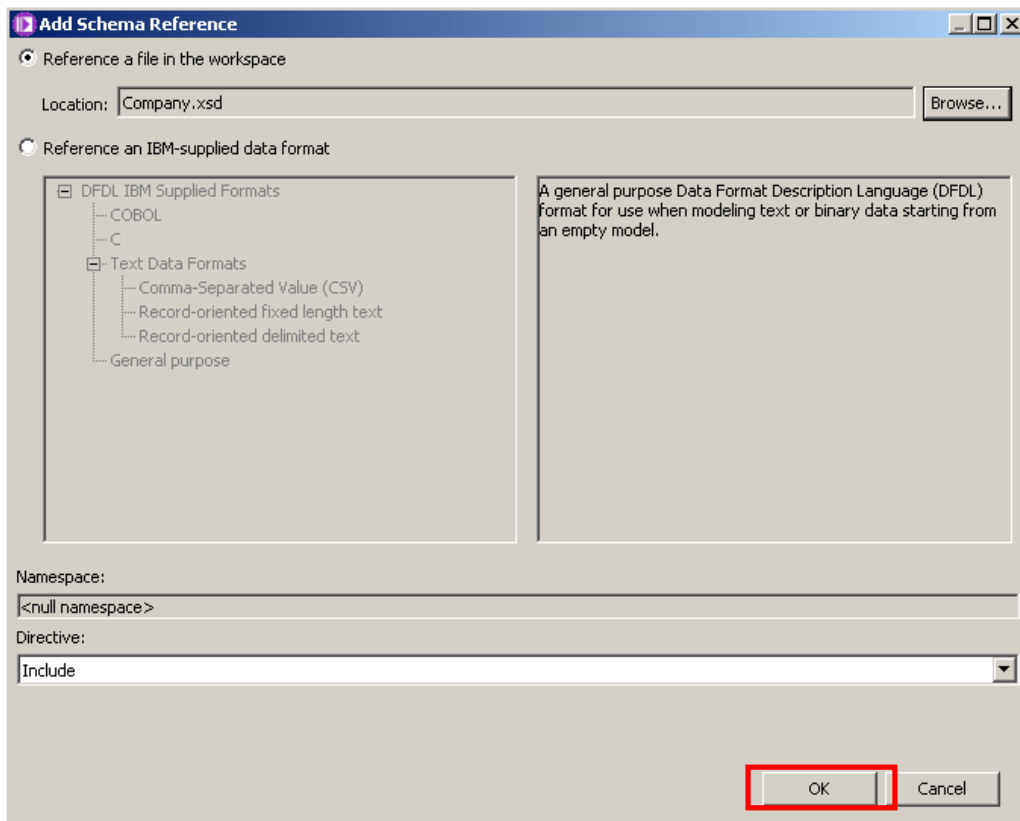


5. Select the "Company.xsd" file you created in the previous (basic) lab.

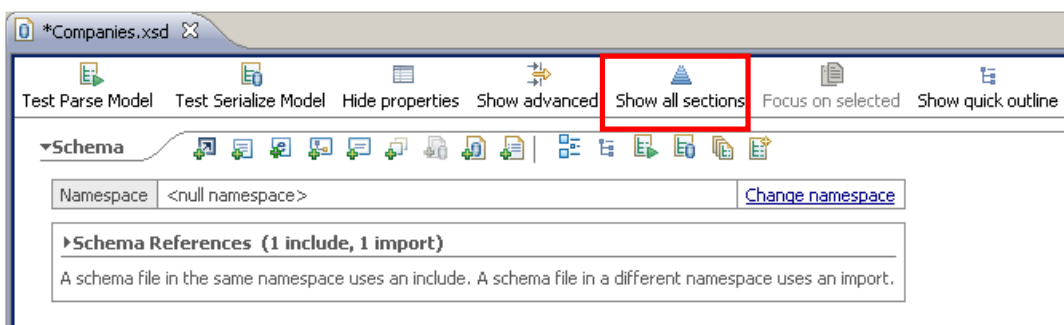
Click OK



- Back in the "Add Schema Reference" window click OK.

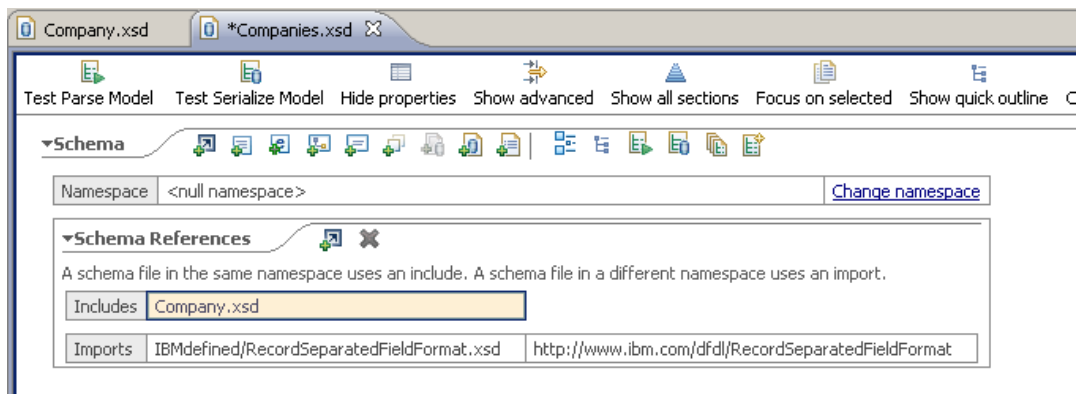


- Click on the "Show all sections" icon to see the whole schema file.

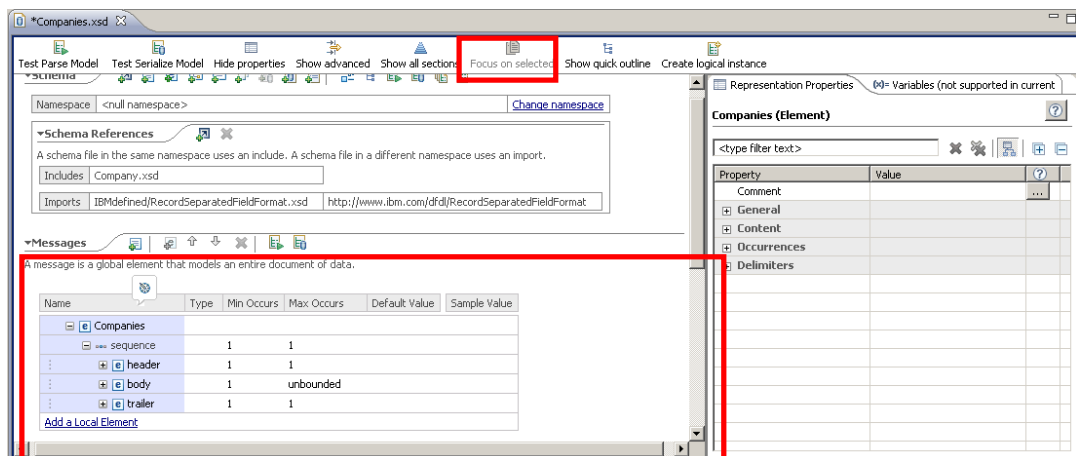


8. Note that you have two Schema References:

1. Company.xsd (which you've just added): you are going to reuse this message model to create a more complex one.
2. RecordSeparatedFieldFormat.xsd (which was automatically added by the wizard): It contains Record Separated specific defaults for DFDL properties.



9. Click anywhere inside the Messages section, and click on the "Focus on selected" icon.



10. Expand the header element by clicking on the "+".

Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
[-] e Companies					
[-] sequence		1	1		
⊕ e header		1	1		
⊕ e body		1	unbounded		
⊕ e trailer		1	1		
Add a Local Element					

11. Change the header element's name to "Header" (capital "H") by clicking on it, and overtyping.

Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
[-] [e] Companies					
[-] sequence		1	1		
[e] Header		1	1		
[-] sequence		1	1		
[e] head_elem1	string	1	1		head_value1
[e] head_elem2	string	1	1		head_value2
[+] [e] body		1	unbounded		
[+] [e] trailer		1	1		

[Add a Local Element](#)

12. Change the 2 elements under Header to "RecordDescription" and "CompanyCount".

Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
[-] [e] Companies					
[-] sequence		1	1		
[e] Header		1	1		
[-] sequence		1	1		
[e] RecordDescription	string	1	1		head_value1
[e] CompanyCount	string	1	1		head_value2
[+] [e] body		1	unbounded		
[+] [e] trailer		1	1		

[Add a Local Element](#)

13. Delete the "body" element by right-clicking on the line of the element and selecting Delete.
(Do not right-click on the text of the element name ... you will see a different context menu).

Name	Type	Min Occurs	Max Occurs
[-] [e] Companies			
[-] ... sequence		1	1
[e] Header		1	1
[-] ... sequence		1	1
[e] RecordDescription	string	1	1
[e] CompanyCount	string	1	1
[+] [e] body		1	unbounded
[e] ...			1

Make Local Element Global	Alt+Shift+E
Move to a New Model Group...	Alt+Shift+G
Move Up	Alt+Up
Move Down	Alt+Down
Cut	Ctrl+X
Copy	Ctrl+C
Paste	Ctrl+V
Delete	Delete
Add Sequence	Ctrl+L, S
Add Choice	Ctrl+L, C

14. Change the trailer element's name to "Trailer" (capital "T").

Change the name of the element under Trailer to "chksum".

Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
[-] Companies					
[-] sequence		1	1		
[-] Header		1	1		
[-] sequence		1	1		
[-] RecordDescription	string	1	1		head_value1
[-] CompanyCount	string	1	1		head_value2
[-] Trailer		1	1		
[-] sequence		1	1		
[-] chksum	string	1	1		trailer_value1

[Add a Local Element](#)

15. Now you are going to define a reference to the Company.xsd model. To do this, you must first create a new sequence element in the Companies model. Right-click on the Companies element and select "Add Sequence".

Name	Type	Min Occurs	Max Occurs	Default Value
[-] Companies				
[-] sequence		1	1	
[-] Header		1	1	
[-] sequence		1	1	
[-] RecordDescription	string	1	1	
[-] CompanyCount	string	1	1	
[-] Trailer		1	1	
[-] sequence		1	1	
[-] chksum	string	1	1	

[Ad](#)

Context menu options:

- Paste (Ctrl+V)
- Delete (Delete)
- Add Sequence (Ctrl+L, S)**
- Add Choice (Ctrl+L, C)
- Test Parse Model (Ctrl+T, P)
- Test Serialize Model (Ctrl+T, S)
- Create a Logical Instance From Model (Ctrl+T, L)
- Show In

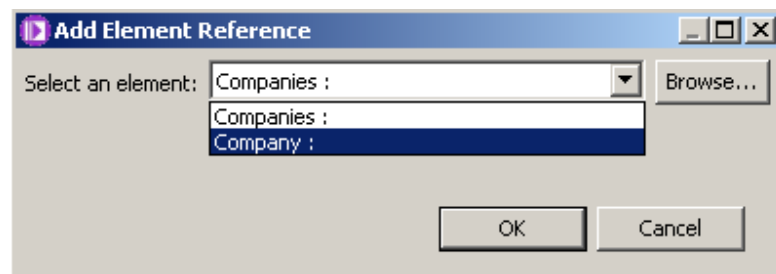
16. Right-click the new sequence element and select “Add Element Reference”.

Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
[-] [e] Companies					
[-] [e] sequence		1	1		
[e] Header		1	1		
[-] [e] sequence		1	1		
[e] RecordDescription	string	1	1		head_value
[e] CompanyCount	string	1	1		head_value
[-] [e] Trailer		1	1		
[-] [e] sequence		1	1		
[e] checksum	string	1	1		trailer_value
[+] [e] sequence		1	1		

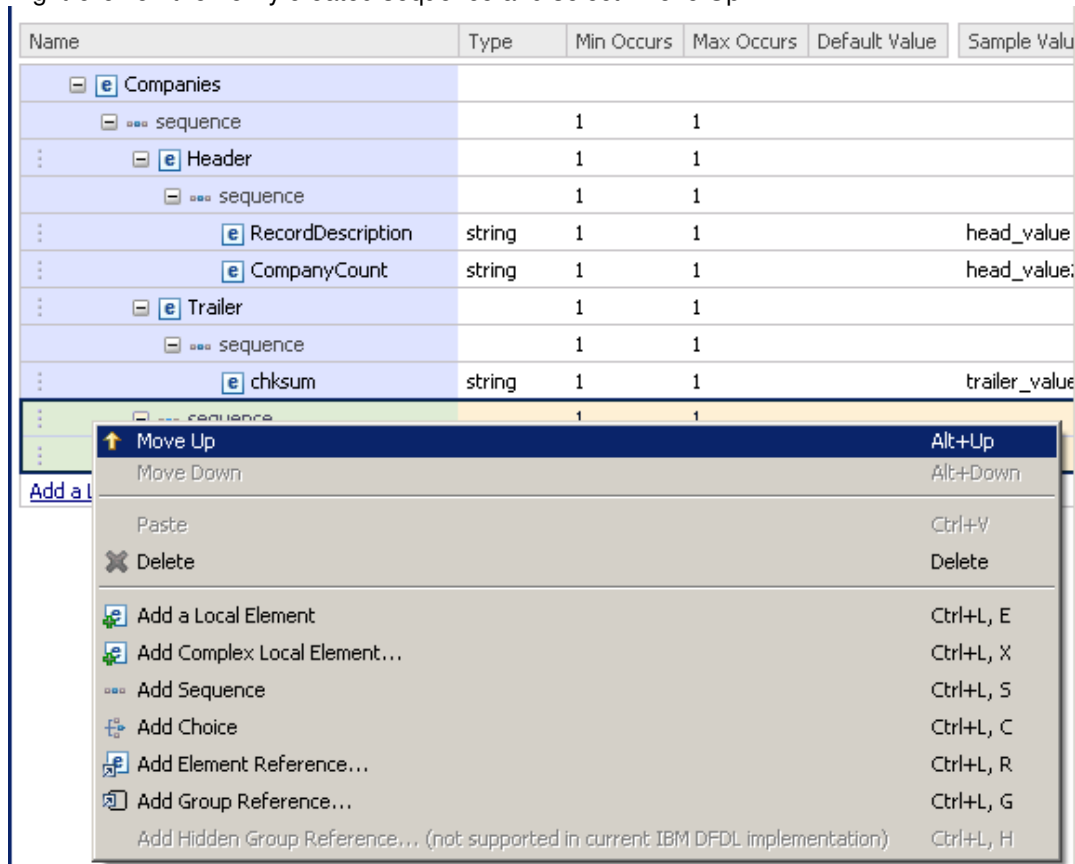
Move Up	Alt+Up
Move Down	Alt+Down
Paste	Ctrl+V
Delete	Delete
Add a Local Element	Ctrl+L, E
Add Complex Local Element...	Ctrl+L, X
Add Sequence	Ctrl+L, S
Add Choice	Ctrl+L, C
Add Element Reference...	Ctrl+L, R
Add Group Reference...	Ctrl+L, G
Add Hidden Group Reference... (not supported in current IBM DFDL implementation)	Ctrl+L, H

17. Select "Company :" from the dropdown

Click OK.

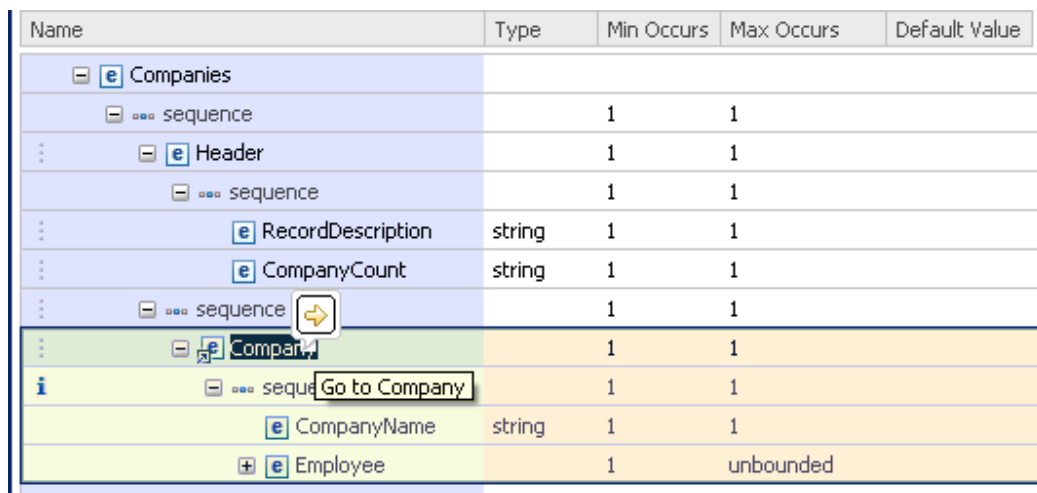


18. Right-click on the newly created sequence and select "Move Up".



Note that the added element reference has a different icon to differentiate from a regular element (this is only a reference to an existing element in other DFDL schema).

Also note that the element reference's name is greyed out because it is read-only. To modify it you will need to open the DFDL schema where it was defined clicking on the yellow arrow that appears when you hover over the element. (→)



- Finally with the new sequence, you have to make an adjustment to the Delimiters Separator property. Highlight the new sequence, and remove the default value of the Separator property. Click Return to make sure the new value is properly recognized by the editor.

The screenshot shows the Message Modeler interface. On the left, a tree view shows a message structure with elements like 'Companies', 'Header', and 'Company'. On the right, the 'Properties' panel is open, showing various settings. The 'Delimiters' section is expanded, and the 'Separator' property is highlighted with the value '<no separator>'.

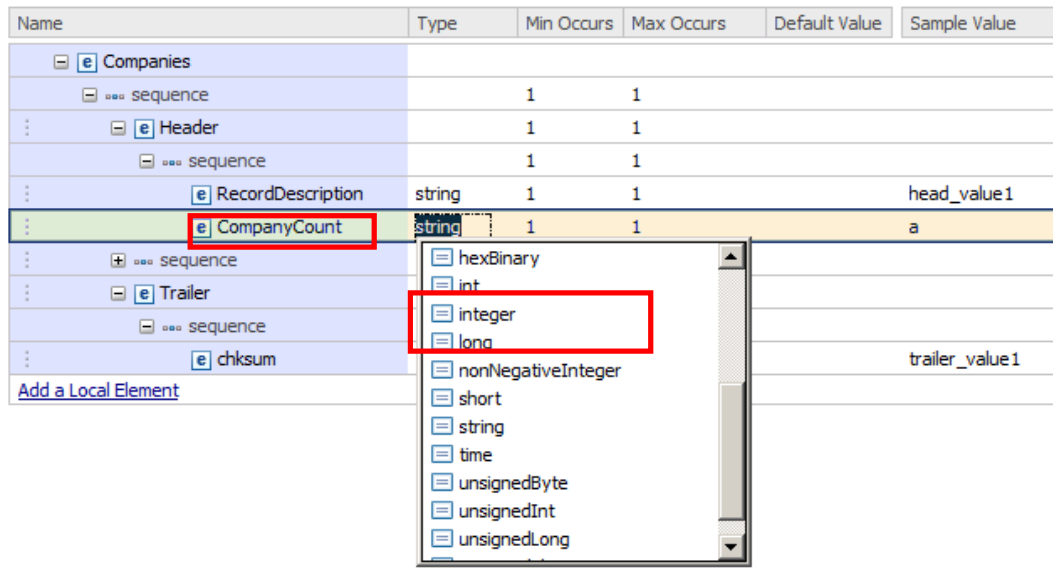
Property	Value
General	
Data Format Reference	<default format>
Encoding (code page)	<dynamically set>
Byte Order	<dynamically set>
Ignore Case	no
Fill Byte	0
Content	
Alignment	
Delimiters	
Separator	<no separator>
Initiator	<no initiator>
Terminator	<no terminator>
Output New Line	%CR;%LF;

- Expand both the Header and Trailer elements by clicking on the "+" next to them.

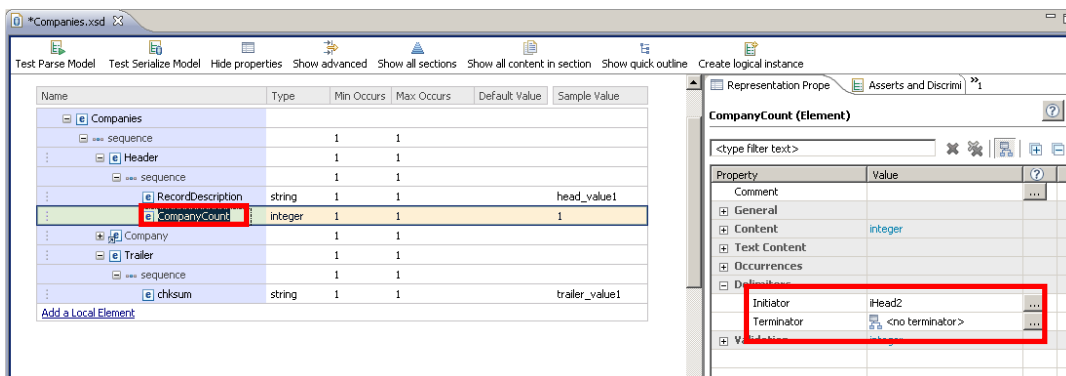
Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
[-] Companies					
[-] sequence		1	1		
[+] Header		1	1		
[+] sequence		1	1		
[+] Trailer		1	1		

[Add a Local Element](#)

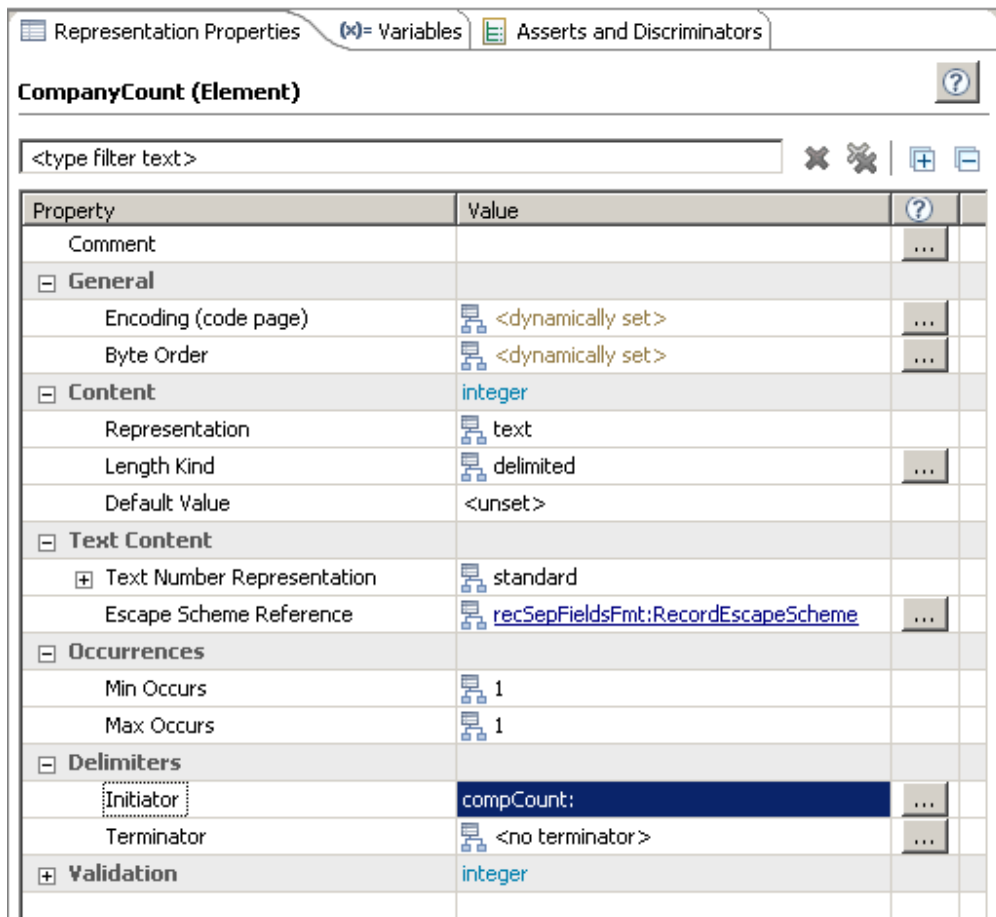
- Change the CompanyCount element's type to "integer" by clicking on its type column and selecting "integer".



- Click on the CompanyCount element again, and look at the "Delimiters" section in the Representation Properties.



23. Change the Initiator's value ("iHead2") to "compCount:" (don't miss out the : (colon)).



The screenshot shows the 'CompanyCount (Element)' properties window. The 'Delimiters' section is expanded, and the 'Initiator' property is selected. The value of the Initiator is 'compCount:'. The 'Terminator' property is set to '<no terminator>'. The 'Validation' property is set to 'integer'.

Property	Value	
Comment		...
General		
Encoding (code page)	<dynamically set>	...
Byte Order	<dynamically set>	...
Content	integer	
Representation	text	
Length Kind	delimited	...
Default Value	<unset>	
Text Content		
Text Number Representation	standard	
Escape Scheme Reference	recSepFieldsFmt:RecordEscapeScheme	...
Occurrences		
Min Occurs	1	
Max Occurs	1	
Delimiters		
Initiator	compCount:	...
Terminator	<no terminator>	...
Validation	integer	

24. Repeat the previous step for the "RecordDescription" and chksum fields with the following values:

RecordDescription	recDesc:
chksum	chksum:

Again, do not miss the colon characters.

Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
Companies	sequence	1	1		
Header	sequence	1	1		
RecordDescription	string	1	1		head_value1
CompanyCount	integer	1	1		1
Company	sequence	1	1		
Trailer	sequence	1	1		
chksum	string	1	1		trailer_value1

Property	Value
Comment	
General	
Content	string
Text Content	
Occurrences	
Delimiters	
Initiator	recDesc:
Terminator	<no terminator>
Validation	string

Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
Companies	sequence	1	1		
Header	sequence	1	1		
RecordDescription	string	1	1		head_value1
CompanyCount	integer	1	1		1
Company	sequence	1	1		
Trailer	sequence	1	1		
chksum	string	1	1		trailer_value1

Property	Value
Comment	
General	
Content	string
Text Content	
Occurrences	
Delimiters	
Initiator	chksum:
Terminator	<no terminator>
Validation	string

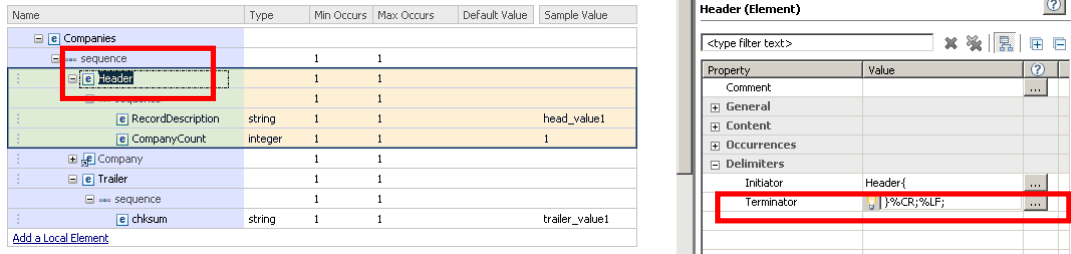
25. Now click on the <sequence> content of the Companies element. In the Delimiters section of the Representation Properties, delete the value of the Separator property. Click Return to make sure the value is updated.

(this is a separator automatically added by the wizard, which is not needed in this case)

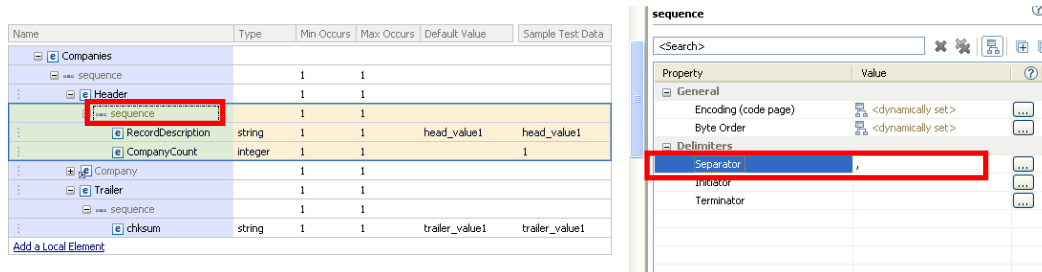
Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
Companies	sequence	1	1		
Header	sequence	1	1		
RecordDescription	string	1	1		head_value1
CompanyCount	integer	1	1		1
Company	sequence	1	1		
Trailer	sequence	1	1		
chksum	string	1	1		trailer_value1

Property	Value
General	
Delimiters	
Separator	
Initiator	<no initiator>
Terminator	<no terminator>

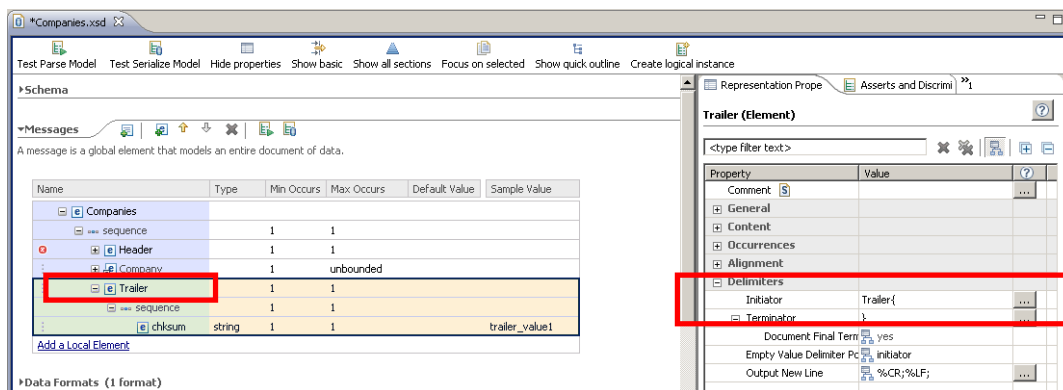
- Click on the Header element. In the Delimiters section of the Representation Properties, set the Terminator property to "}%CR;%LF;"



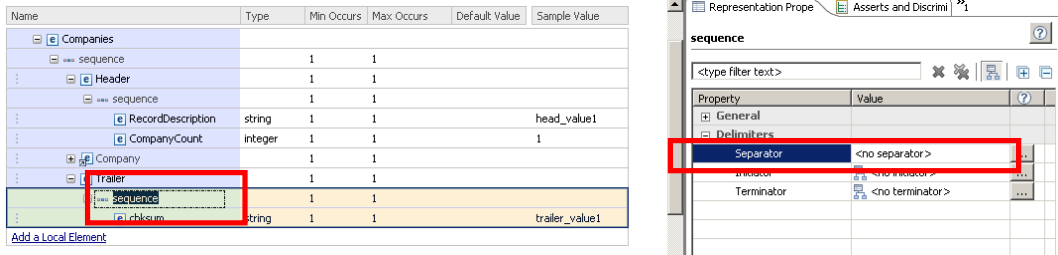
- Then click on the <sequence> content of the Header element and in the Delimiters section of the Representation Properties, set the Separator property to "," (comma)



- Click on the Trailer element, and in the Delimiters section of the Representation Properties, set the Terminator property to "}".



- Then click on the <sequence> content of the Trailer element and in the Delimiters section of the Representation Properties, delete the Separator property's value.

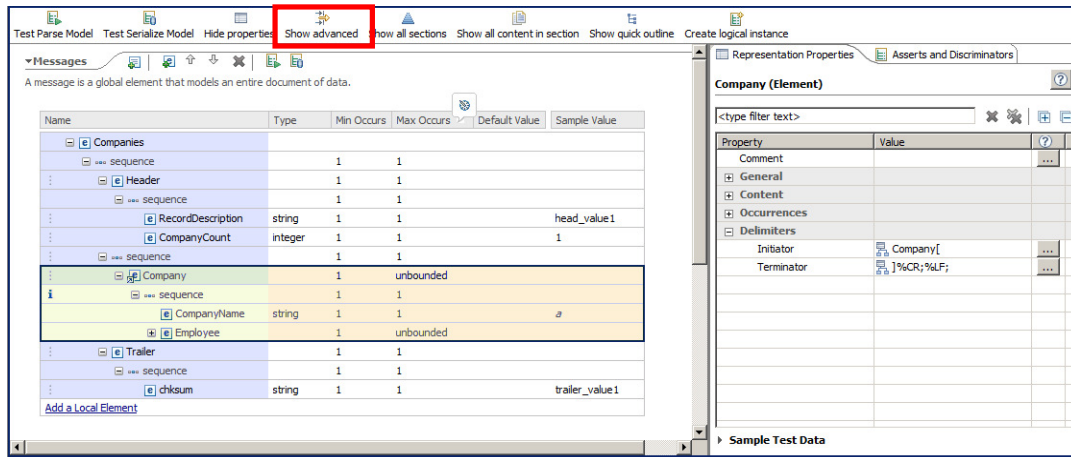


- Click on the "Max Occurs" column of the Company element reference, and change it from "1" to "unbounded".

This will allow the Company element reference to have infinite occurrences.

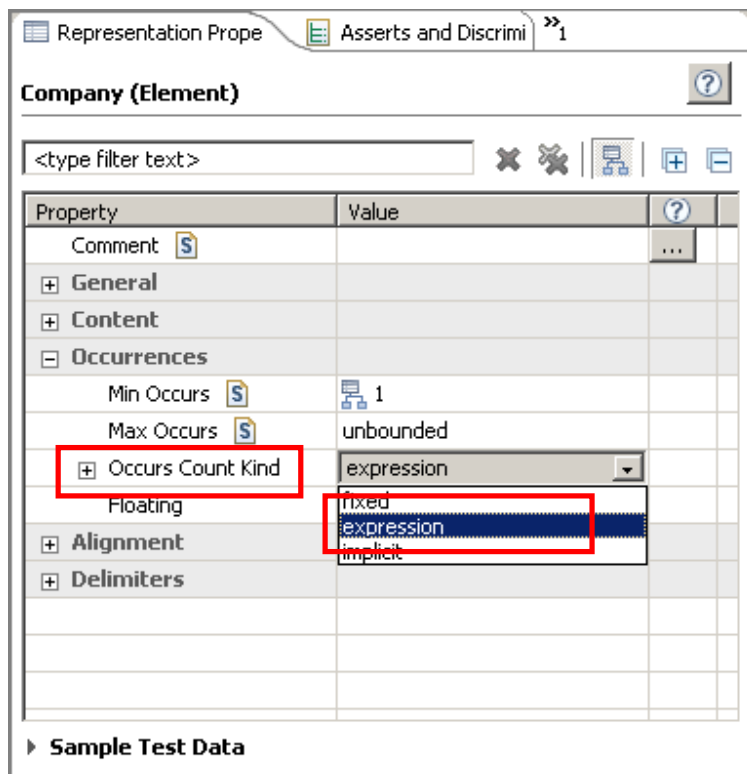
Name	Type	Min Occurs	Max Occurs	Default Value	Sample Value
Companies					
sequence		1	1		
Header		1	1		
sequence		1	1		
RecordDescription	string	1	1	head_value1	
CompanyCount	integer	1	1	1	
sequence		1	1		
Company		1	unbounded		
sequence		1	unbounded		
CompanyName	string	1	1		
Employee		1	unbounded		
Trailer		1	1		
sequence		1	1		
chksum	string	1	1		trailer_value1

- Click on the "Show Advanced" icon to show the advanced Representation Properties.



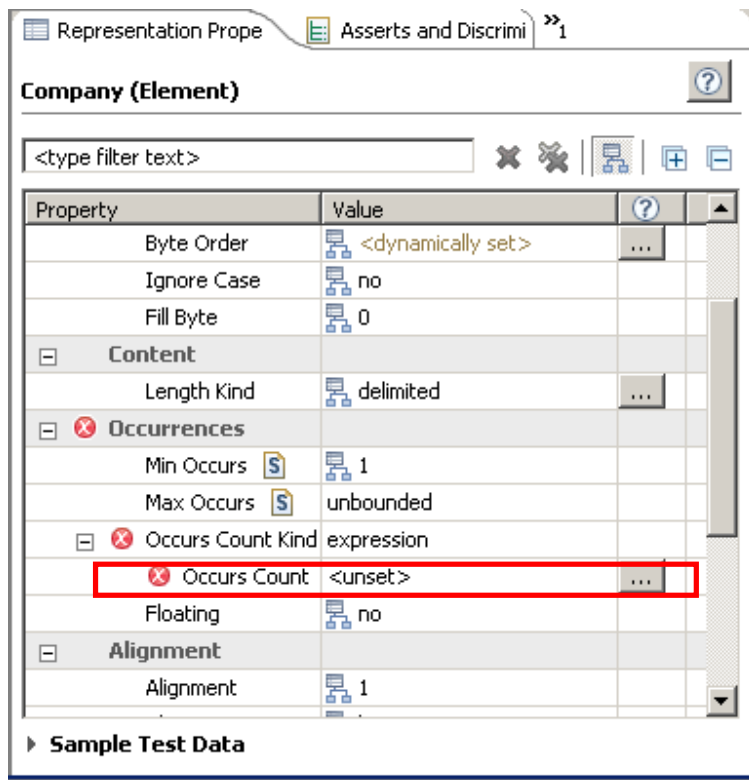
- Click on the Company element reference, and in the Occurrences section of the Representation Properties, change the "Occurs Count Kind" from "fixed" to "expression".

You may need to Save the model and close and reopen the schema to update the Representation Properties, so that the "Occurs Count Kind" property appears.



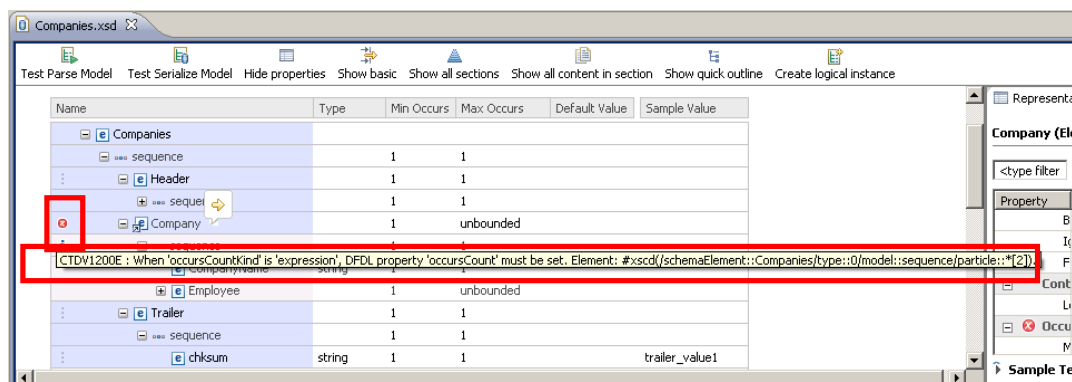
33. Expand the "Occurs Count Kind" property to show the "Occurs Count" property.

Leave the field empty and save your XSD by pressing Ctrl+S or File->S.



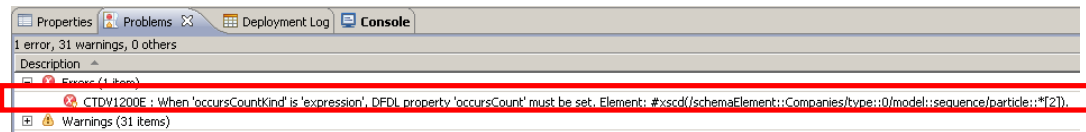
When saving, the DFDL Editor also validates the schema file.

34. Note that an error icon has appeared next to the Company element and 'Occurrences'.

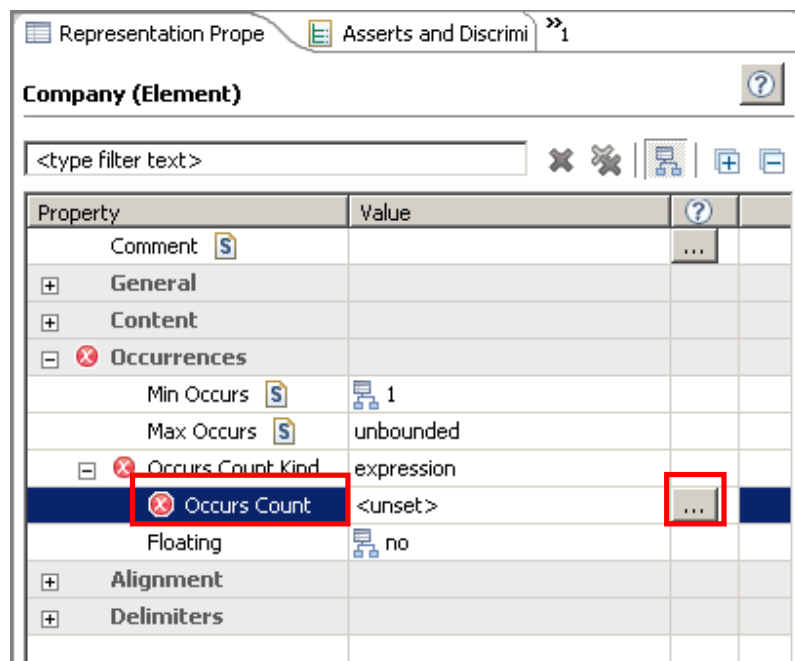


Hover over the icon to display the error description. It states that the OccursCount property can't be empty.

35. Click on the Problems tab.



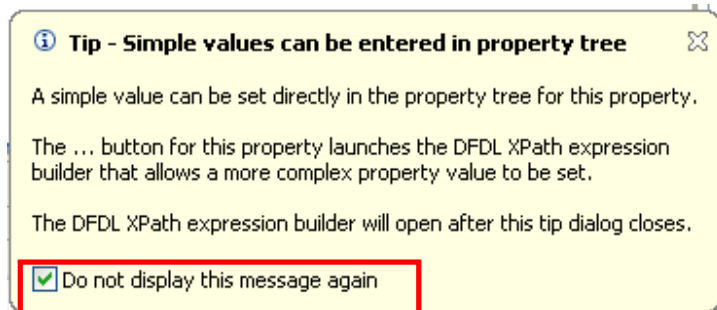
Double-click on the error and the representation property of the problematic element will open.



Click on the button next to the OccursCount property.

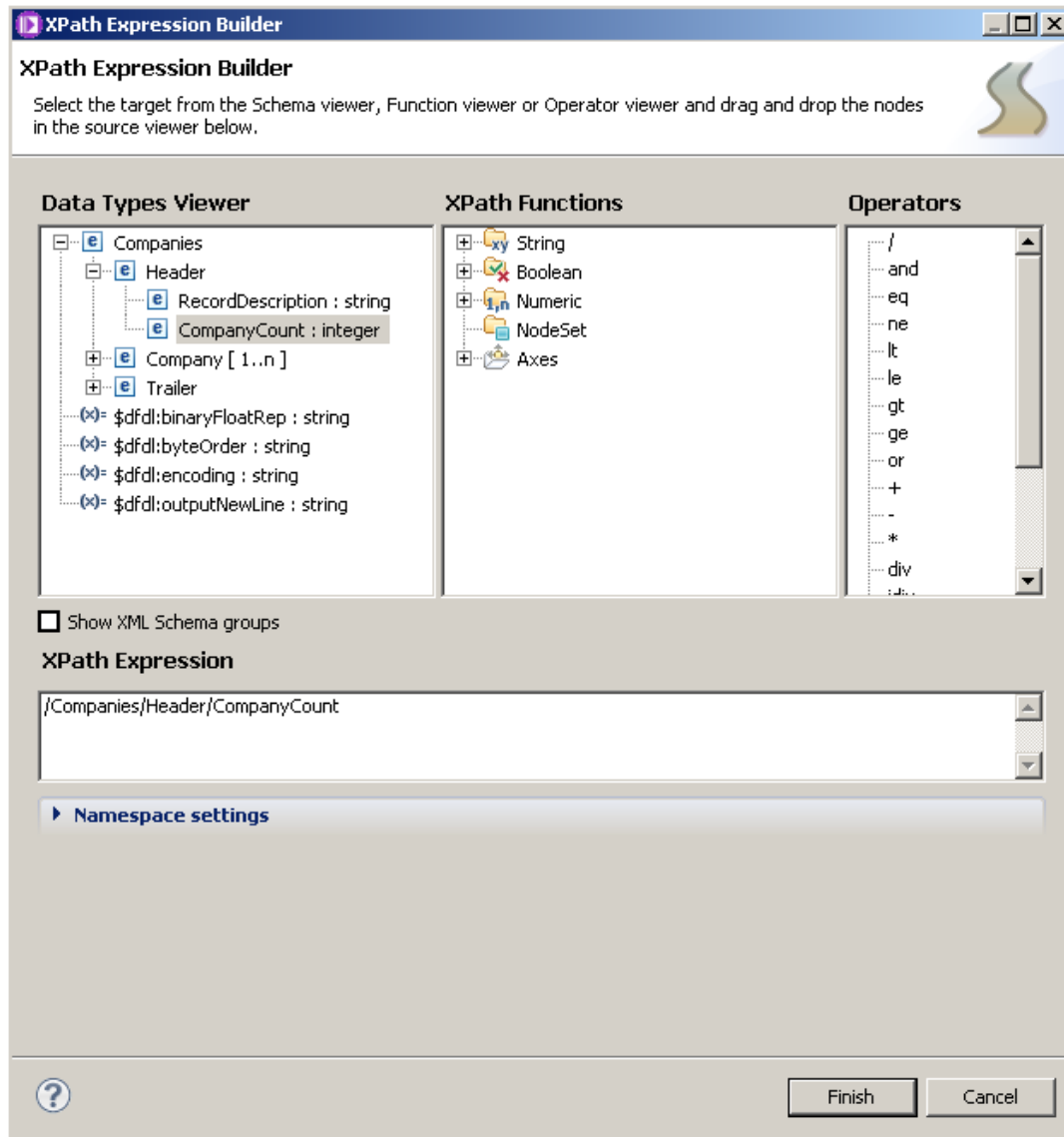
You may need to close and reopen the model to ensure the "Occurs Count" property appears.

36. A message bubble will appear. Check "Do not display this message again" and close it by clicking on the "X".



37. In the XPath Expression Builder window expand the Header element and double-click on the CompanyCount element.

Click on the Finish button.

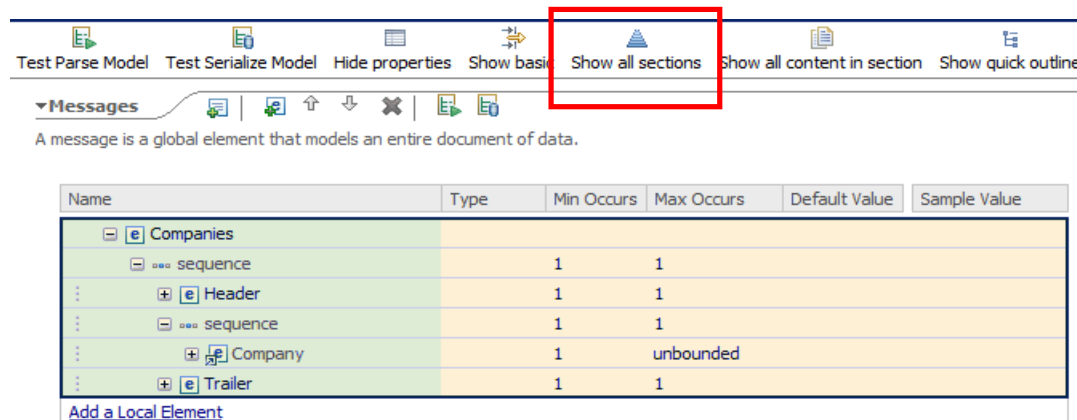


This will indicate that the number of occurrences of the Company element is dictated by the CompanyCount element in the Header.

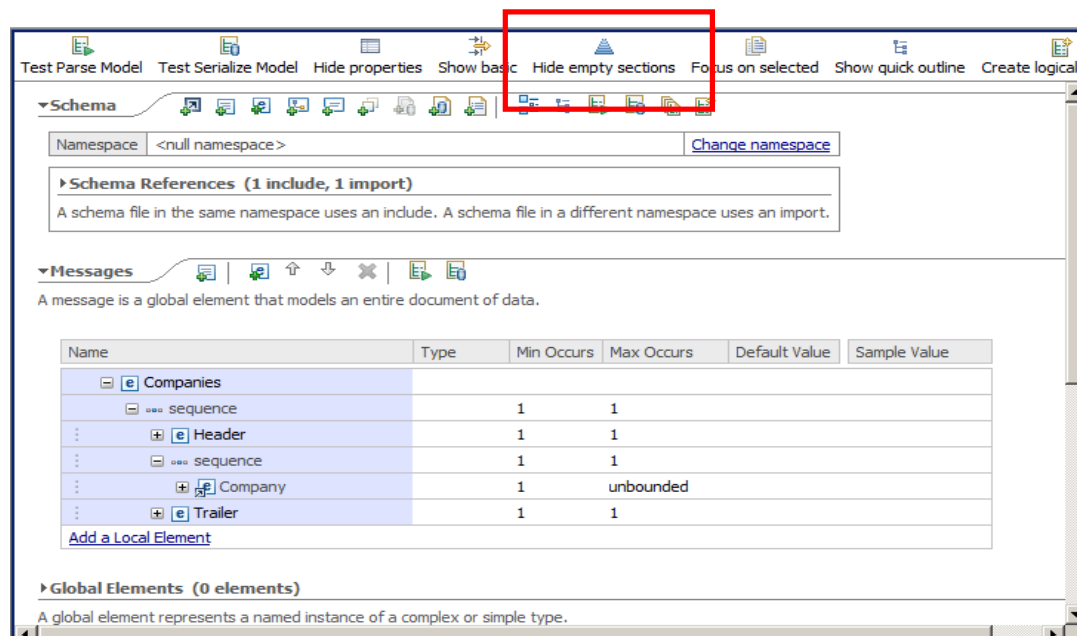
- 38. Finally, the input file may, or may not, have a Carriage Return Line Feed as the last character. In some cases, the last record may be missing the final “new line” character.

To handle this situation, we will change the default values for this model so that it will be able to handle both scenarios (ie. it will parse successfully, irrespective of whether the final character is present or not).

In the editor, collapse the Companies model, and then click “Show all sections”.



- 39. To make the display a little less busy, click “Hide empty sections”.



40. Expand the Data Formats section.

Highlight the <default format> field. This is where you can define many default values for the message model.

In the Representation Properties, expand the Delimiters section, and locate the property DocumentFinalTerminatorCanBeMissing. Set this property to “yes”. Ensure you press the Return key to ensure the property is correctly updated.

The screenshot shows the IBM Integration Bus V9.0 Workshop interface. The left pane displays the Schema view with the 'Data Formats' section expanded, showing a table with one entry: '<default format>' of type 'Definition Format'. The right pane shows the 'Representation Properties' for the selected '<default format> (Data Format)'. The 'Delimiters' section is expanded, and the property 'Document Final Terminator Can Be Missing' is set to 'yes'.

Property	Value
Comment	S
General	
Content	
Text Content	
Binary Content	
Structural Content	
Occurrences	
Alignment	
Delimiters	
Separator	,
Separator Policy	suppressed
Separator Position	infix
Initiator	<no initiator>
Terminator	<no terminator>
Document Final Terminator Can Be Missing	yes
Nil Value Delimiter Policy	initiator
Empty Value Delimiter Policy	initiator
Output New Line	%CR;%LF;
Calculated Values	

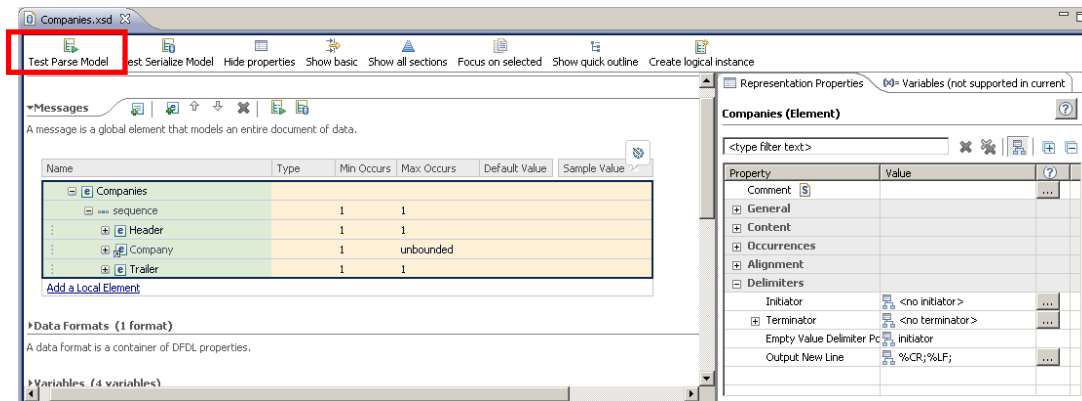
41. Press Ctrl+S or File->Save to save your work.

Check that the error has gone and the Problems view is clean.

5. Testing the Message Model

1. Now that the Message Model is complete, you will test it against a delimited file.

Click on the Test Parse Model icon.



2. In the Parser Input section, select "Content from a data file" and click the Browse button.

Test Parse Model

Message
Select message for testing. [More...](#)
Message name:* Companies

Parser Input
Select content to be parsed against schema.
 Content from 'DFDL Test - Serialize' view
 Content from a data file
Input file name:*

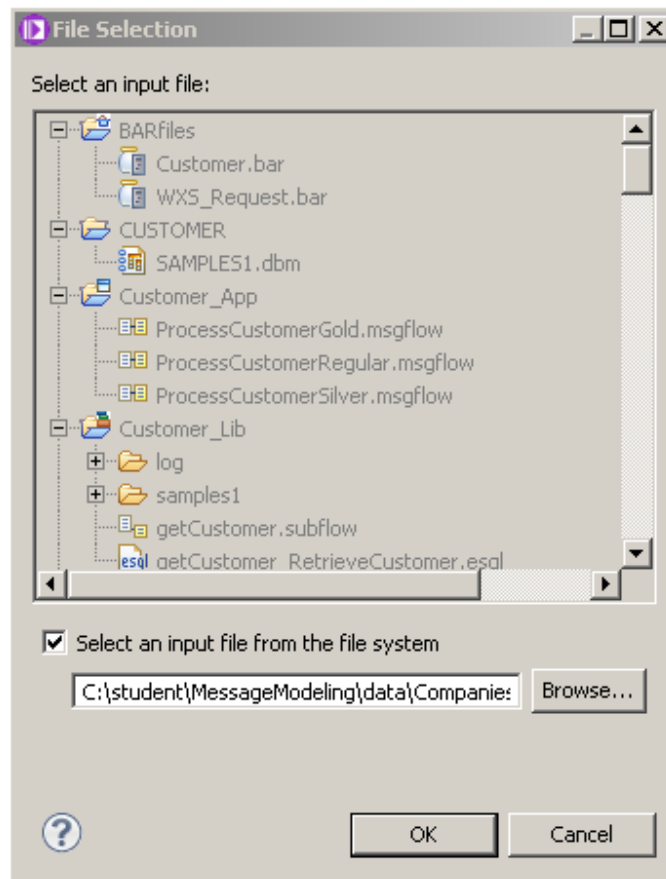
Specify runtime configuration.

Runtime encoding options
Provide runtime values for properties which have been configured in the model to be dynamically set. [More...](#)
Encoding (code page): UTF-8
Floating point format: IEEE Non-Extended
Byte order: Little endian Big endian

Runtime validation
 Validate data against schema [More...](#)

3. Check the "Select an input file from the file system" and browse to "C:\student\MessageModeling\data" and select the "Companies.txt" file.

Click OK.



4. Leave all the default values and click OK.

Test Parse Model

Message
Select message for testing. [More...](#)
Message name: * Companies

Parser Input
Select content to be parsed against schema.
 Content from 'DFDL Test - Serialize' view
 Content from a data file
Input file name: * C:\student\MessageModeling\data\Companies.txt [Browse...](#)

Specify runtime configuration.

Runtime encoding options
Provide runtime values for properties which have been configured in the model to be dynamically set.
[More...](#)
Encoding (code page): UTF-8
Floating point format: IEEE Non-Extended
Byte order: Little endian Big endian

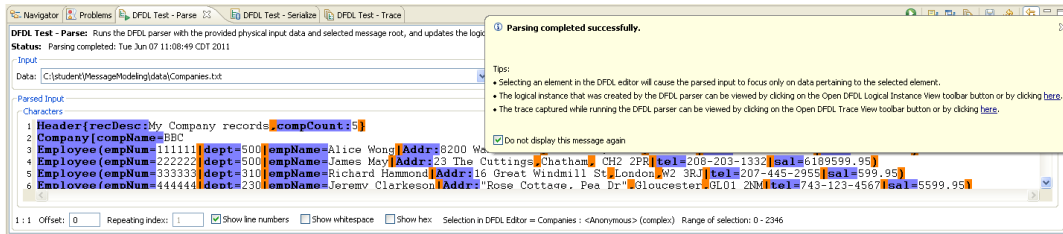
Runtime validation
 Validate data against schema [More...](#)

[Restore Defaults](#)

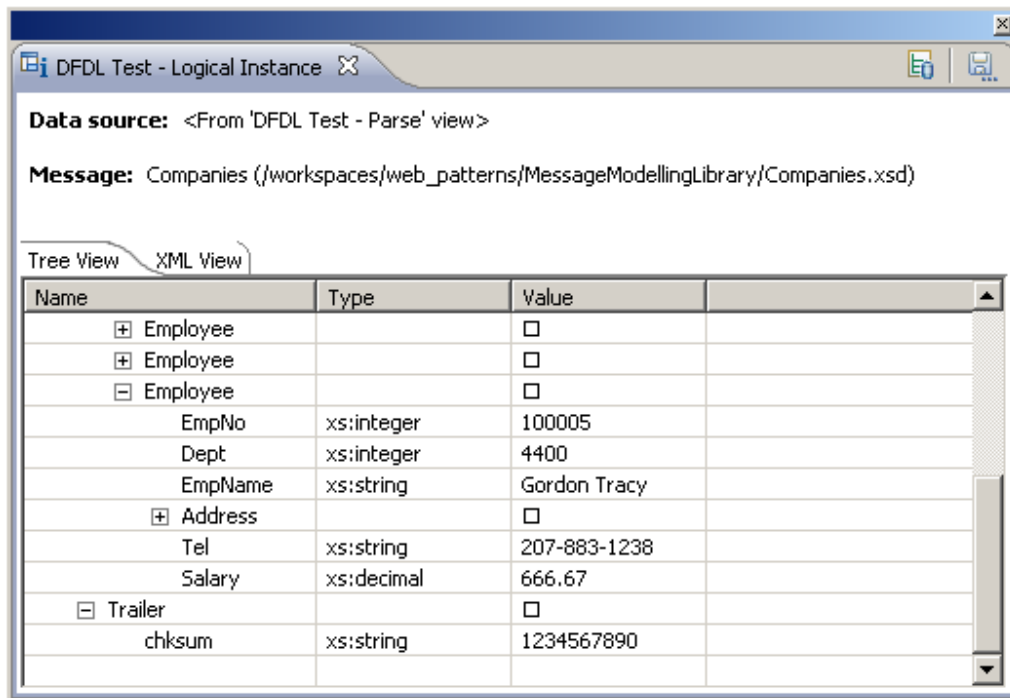
OK Cancel

- A message bubble with a "Parsing completed successfully" should appear.

Check the "Do not display this message again" and close it by clicking on the "X", or click anywhere on the test parse window.



- Review the Parsed input and the Logical Instance view to verify the parsing was correct.



This concludes the Advanced Record-oriented Text Message Modeling lab.