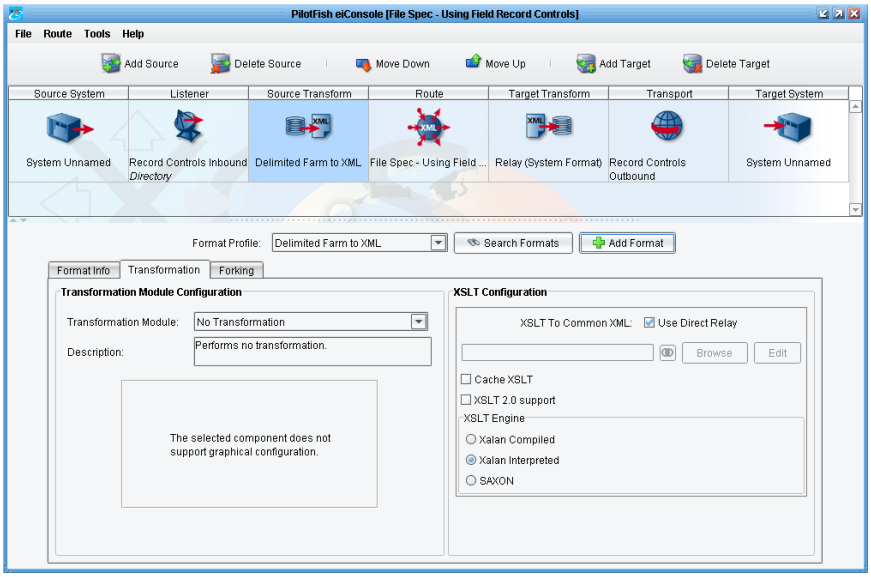
File Specifications – Using Field Record Controls

Overview

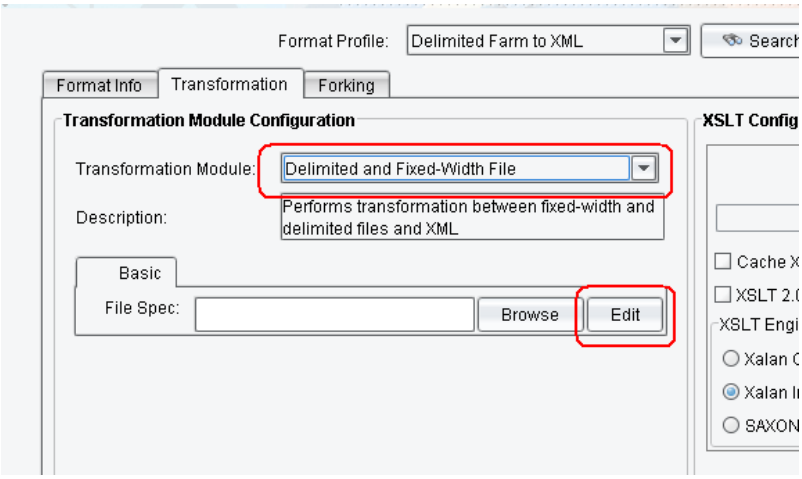
In this tutorial we'll cover using "Field Record Controls" in File Specifications to conditionally select Records in a fixed-width format. This tutorial expands on the concepts of "File Specifications – Handling Fixed-Width Files," so users are expected to be familiar with that content.

Steps

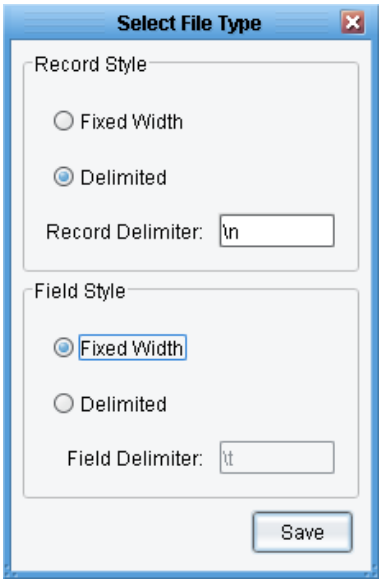
Begin by creating and configuring a new Route similar to the one created in "File Specifications – Handling Fixed-Width Files." Select the "Source Transform" stage and add a new Format, this time called "Delimited Farm to XML":



Select the "Delimited and Fixed-Width File" Transformer and click Edit to open the File Specification Editor:



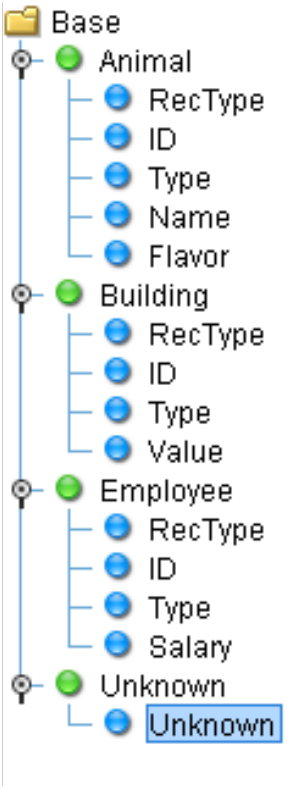
Select "Other" from the drop-down and configure this File Specification with "\n"-delimited Records and Fixed-Width Fields:



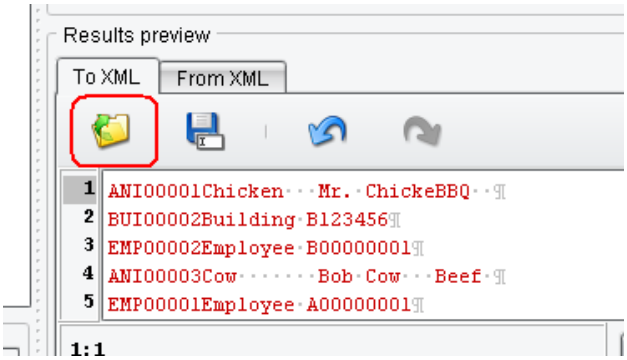
Create the following Records (in green) and add Fields (in blue) with the specified lengths (in [brackets]):

- Animal
  - RecType [3]
  - ID [5]
  - Type [10]
  - Name [10]
  - Flavor [5]
- Building
  - RecType [3]
  - ID [5]
  - Type [10]
  - Value [6]
- Employee
  - RecType [3]
  - ID [5]
  - Type [10]
  - Salary [8]
- Unknown
  - Unknown [10]

When completed, the structure should look like this:



Load in the "farm_fixed.txt" sample file:

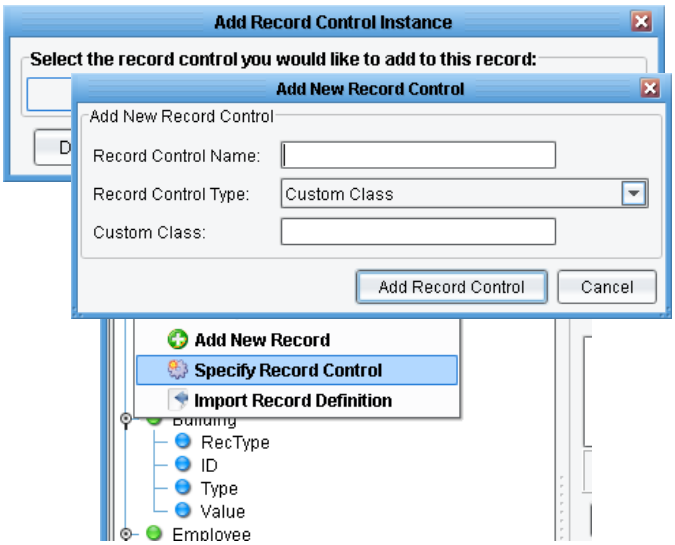


There is one key difference between this File Specification and the one covered in the fixed-width tutorial: there are multiple Records. As you can see from the sample, there is no guaranteed order for the Records, nor is there any indication that there's a particular occurrence limit; each Record can occur any number of times. Another facet is that each Record's first three characters define the Record type. For example, “ANI” indicates an “Animal,” while “BUI” indicates a “Building.”

This is a fairly typical pattern for delimited and fixed-width formats. Some Field within the Record definition indicates the type of Record with varying levels of complexity. Some formats make use of many, many nested, conditionally selected Records.

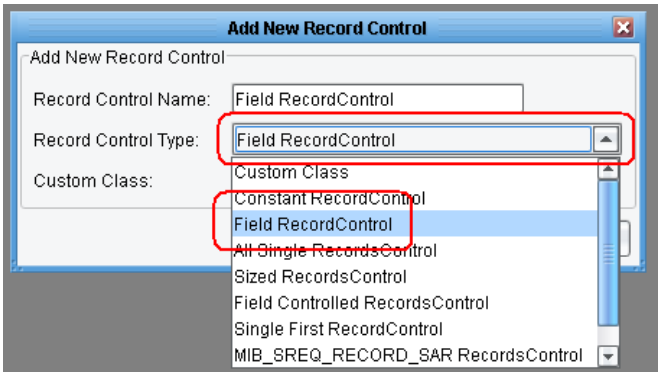
In order to select our correct Record based on the first three characters (the “indicator”), we'll need to make use of a File Specification concept called a “Record Control.” Record Controls are components that select or alter the contents of Records or Fields based on some configuration. The Record Control we'll be using is called the “Field Record Control.” To make use of it, right-click on the “Base” node and select “Specify Record Control”:

This will raise a dialog where you can select from a list of defined Record Controls:

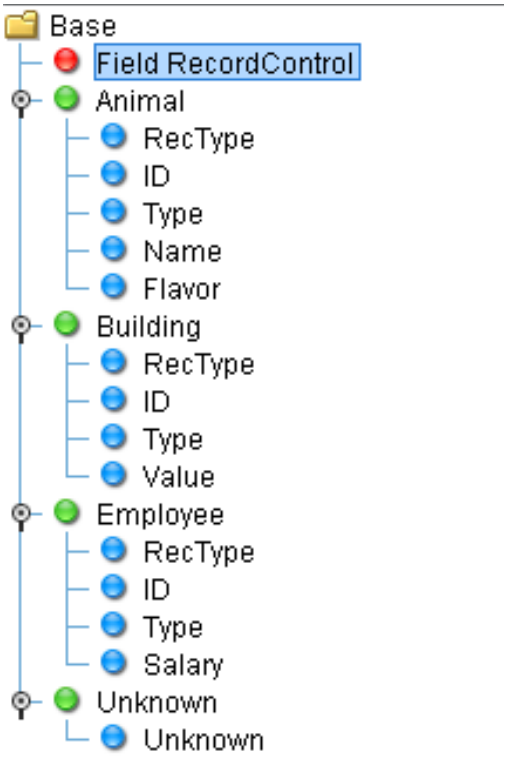


As we have no such Record Controls defined, click the "Define new..." button to raise another dialog:

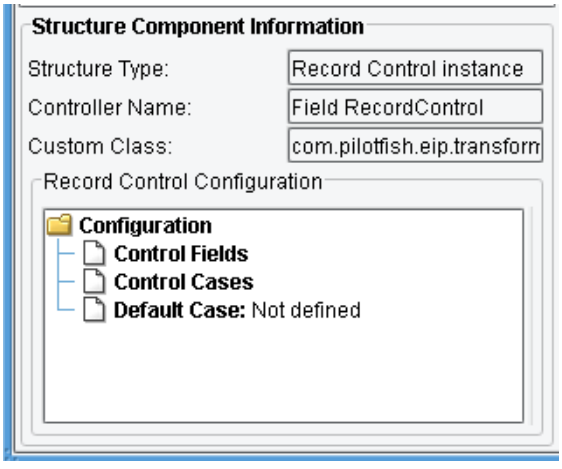
In this dialog, you can specify a "Record Control Type" via drop-down and its name. Select the "Field Record Control" and leave the default name:



Click "Add Record Control," then click "Select." You should now see the Record Control entry in the structure panel:



Select the "Field RecordControl" entry and the bottom configuration panel should change to reflect it:



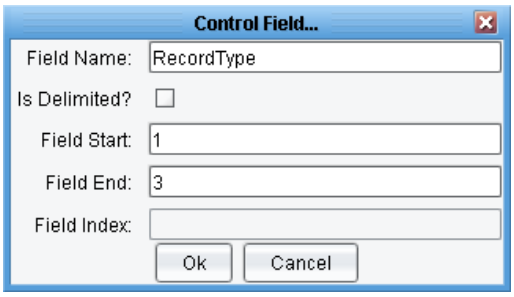
A Field Record Control has three components:

1. Control Fields, which are defined Fields (by position or by start/end) used for testing.
2. Control Cases, which are cases (tests) against Control Fields.
3. Default Cause, which is the default if no Control Cases evaluate to "true".

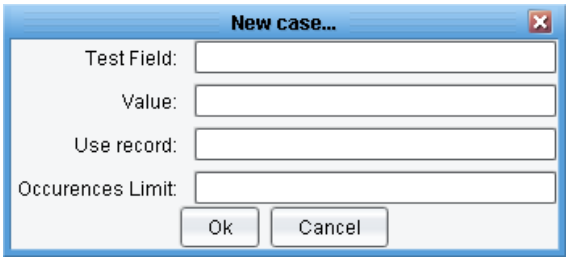
We'll start by defining a Control Field. Right-click the Control Fields entry and select "Add." This will raise a dialog where you can define a Control Field:

The first part is the "Field Name," which is what we'll be using to identify the Control Field in later cases. We'll call ours "RecordType":

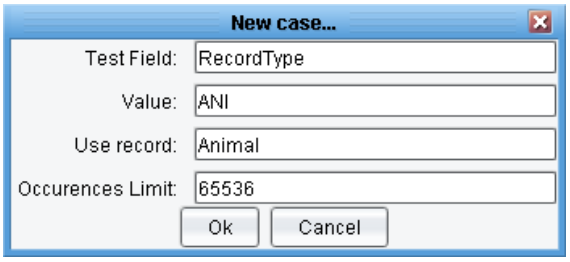
Next, we can define the Start and End of this Control Field. Since we're testing the first three characters for "ANI," "EMP," or "BUI," start from "1" and end at "3":



Click "OK," and then right-click on Control Cases and click "Add" to start defining our tests:

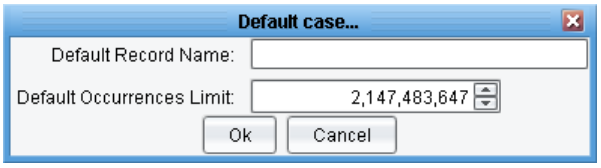


Each Control Case has a Test Field (the name of the Control Field we're testing), a Value (the value we're testing against), a Use Record (which Record to use if the Test Field matches Value), and an Occurrences Limit (how many times this case can apply). We'll provide an entry for "ANI" against the Animal Record:

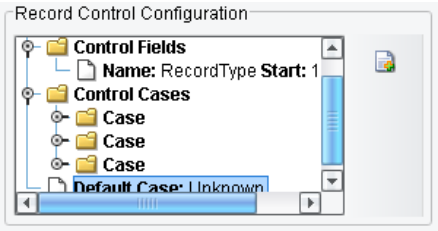


Define similar cases for EMP → Employee and BUI → Building.

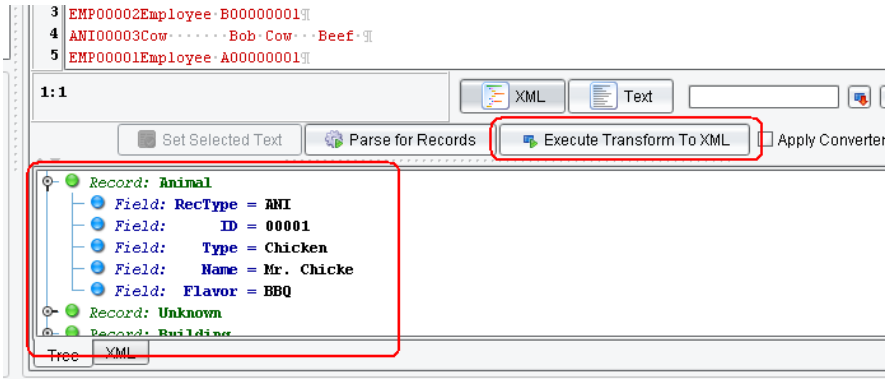
Next, we'll need to define a default Record if no match is found. Right-click on Default Case and select "Define":



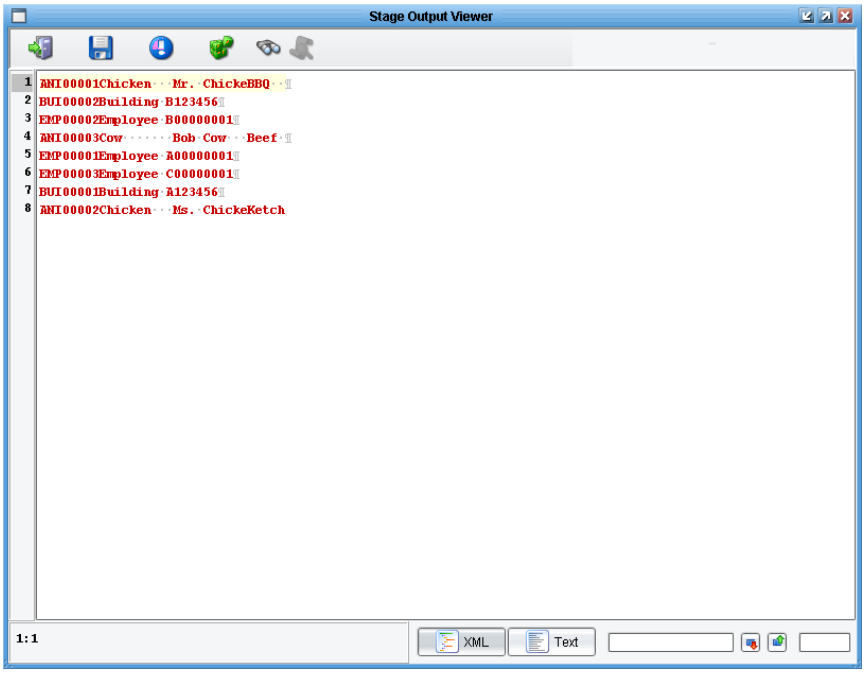
We'll provide "Unknown" for our Default Record Name, then click "OK." Our final configuration should look something like this:



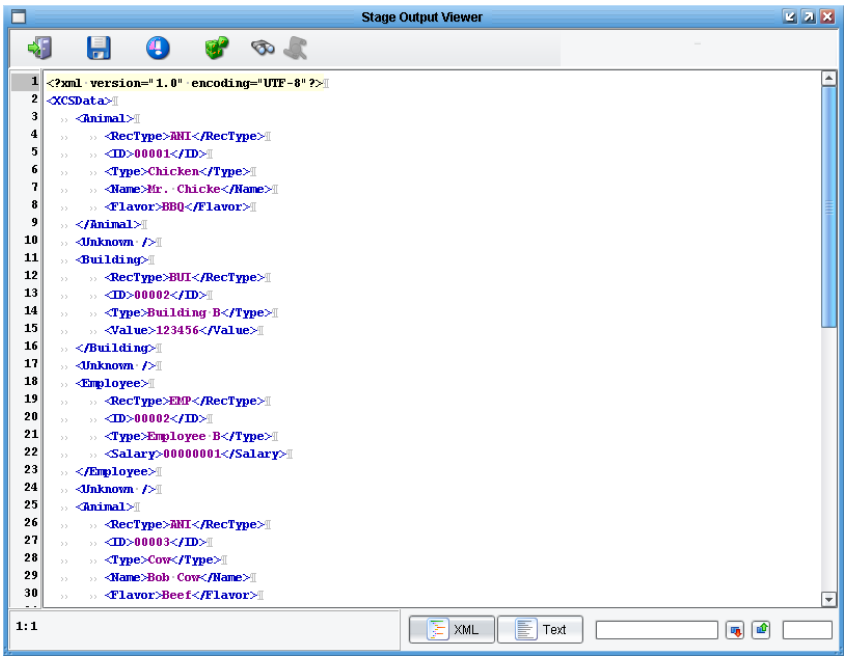
Finally, click the "Execute Transform to XML" button to test the results:



Save the File Specification, close the Editor, return to the eiConsole, and test the Route from end to end using the provided sample file. Here's the resulting "before":



And here's the "after":



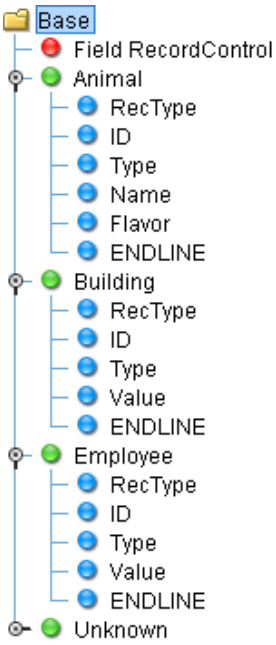
You can use Record Controls inside Records as many layers deep as necessary.

Issues

If you find that you are seeing <Unknown/> entries in your output, this is because some file systems use a 2-character end-line sequence (carriage return + linefeed), while others use a single character.

To fix this problem you should add an extra field to your record definition to consume the additional character. Add an additional record called "ENDLINE" with a length of 1 byte. This should stop the empty records from appearing in your output.

The resulting configuration should look like this:



Which will produce an ouput like this:

Notice that there are no “Unknown” elements present.