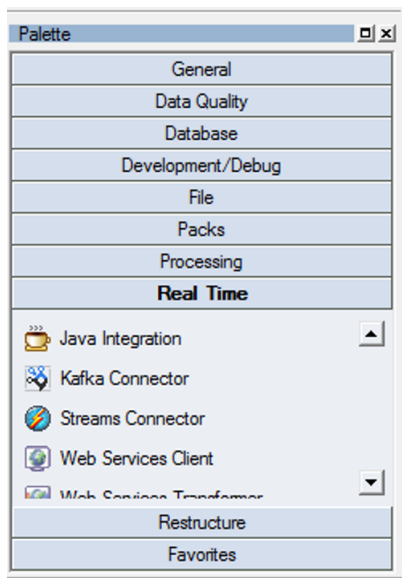


Вывод табличных данных в XML средствами IBM DataStage с дополнительной обработкой в Java-операторе

<https://github.com/zinal/dssamples/tree/main/xml-output/complex>

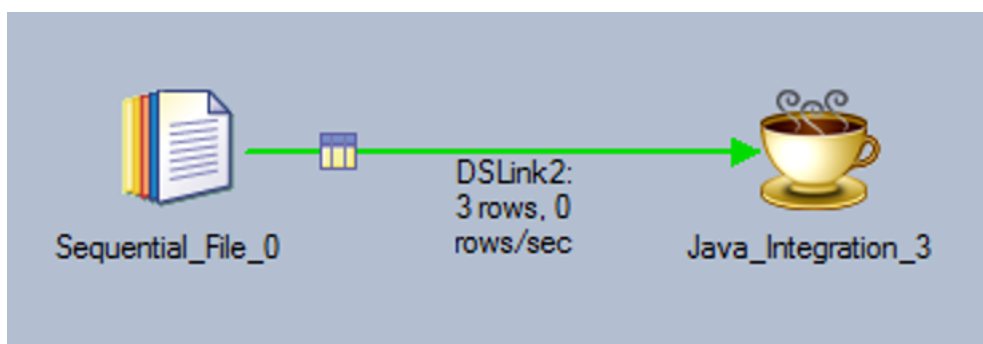
В разделе «Real Time» есть оператор «Java Integration» (см. скриншот ниже).



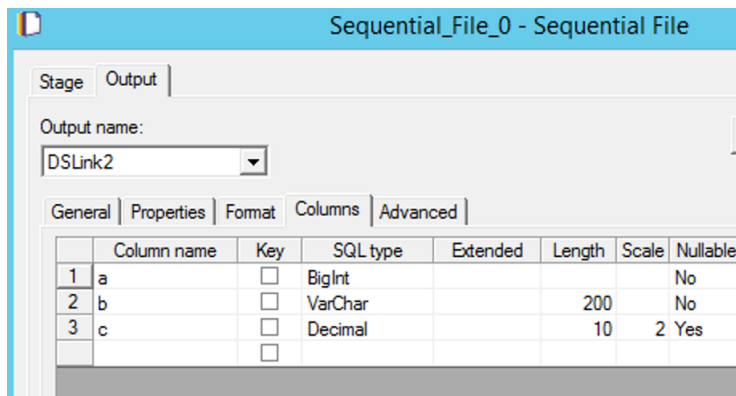
Он позволяет выполнять произвольные виды обработки записей из одного либо нескольких входных потоков, а также возвращать записи в один или несколько выходных потоков.

В нашем простейшем примере у Java-оператора не будет выходных потоков, результаты своей работы он будет записывать в файл.

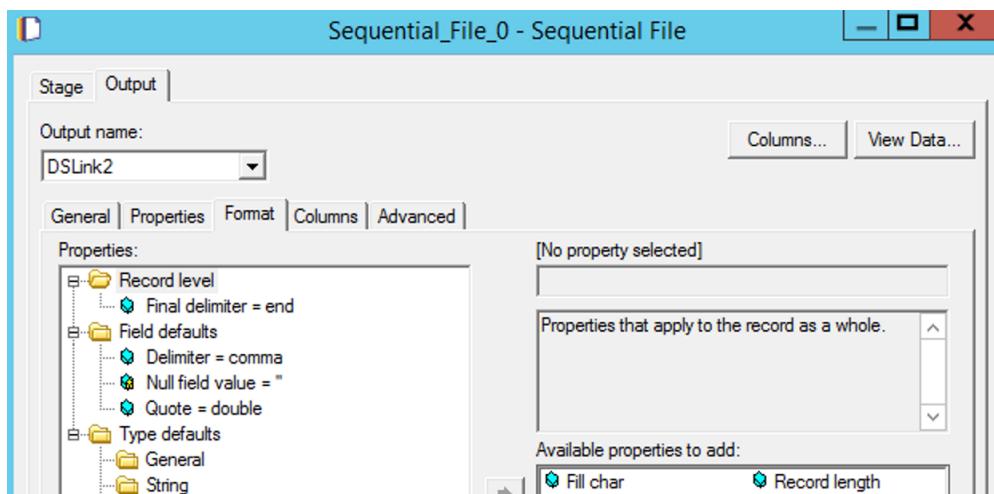
Простейшее задание читает CSV и подаёт на обработку в Java-оператор:



При чтении CSV необходимо определить структуру полей, как показано в примере настроек ниже:



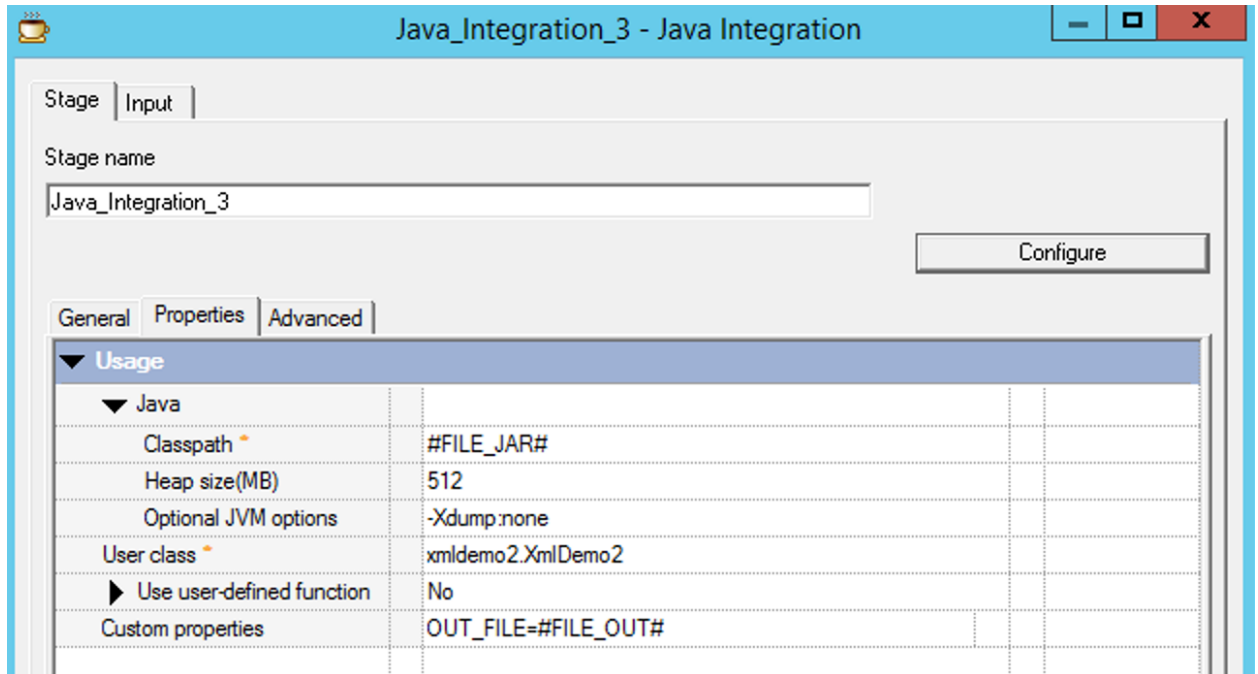
Если предполагаются пустые значения (и их надо обрабатывать как NULL), важно выставить свойство «Null field value» в значение " (две одинарные кавычки – пустое текстовое значение):



Логика вывода данных определяется кодом оператора на языке Java. Например:

```
@Override
public void process() throws Exception {
    final List<ColumnMetadata> meta = inputLink.getColumnMetadata();
    writer.writeStartDocument("utf-8", "1.0");
    writer.writeStartElement("output-data");
    while (true) {
        InputRecord ir = inputLink.readRecord();
        if (ir==null)
            break; // no more input
        writer.writeStartElement("record");
        for (int i=0; i<meta.size(); ++i) {
            Object value = ir.getValue(i);
            if (value!=null) {
                writer.writeStartElement(meta.get(i).getName());
                writer.writeCharacters(value.toString());
                writer.writeEndElement();
            }
        }
        writer.writeEndElement();
    }
    writer.writeEndElement();
    writer.writeEndDocument();
    writer.close();
    writer = null;
    fos.close();
    fos = null;
}
```

В настройках оператора необходимо определить путь к скомпилированному JAR-файлу, имя класса реализации оператора, объём используемой памяти. Рекомендуется также установить параметр «-Xdump:none» для отключения автоматического формирования диагностических дампов при ошибках выполнения Java-оператора:



У Java-оператора могут быть свои собственные настройки, указываемые в разделе «Custom properties». В данном случае имя выходного файла определяется через значение параметра задания DataStage.