

CSVtoJDBC

Overview

Method used to transfer from a CSV file to a database. This database could then be harvested from by using the [JDBCFetch](#) process.

Short Option	Long Option	Parameter Value Map	Description	Required
i	inputFile	FILENAME	csv file to be read into the database	true
d	driver	JDBC_DRIVER	jdbc driver class	true
c	connection	JDBC_CONN	jdbc connection string	true
u	username	USERNAME	database username	true
p	password	PASSWORD	database password	true
t	tableName	TABLE_NAME	a single database table name for the destination	false

Usage

Define Alias

```
CSVtoJDBC="java $OPTS -Dprocess-task=CSVtoJDBC org.vivoweb.harvester.util.CSVtoJDBC"
```

Invocation

```
$CSVtoJDBC -i people.csv -X config/database-connection.xml
```

Here the people.csv is to be put into the database connection described in the database-connection.xml

```
$CSVtoJDBC -i people.csv -d net.sourceforge.jtds.jdbc.Driver -c jdbc:jtds:sqlserver://01.255.01.01/DATA -u USERNAME -p PASSWORD
```

Methods

This class externalizes an already available function of the H2 libraries to turn CSV data into database data. This implementation does put the data into any jdbc database.

Expected CSV format

```
Header1, Header2, Header3, Header4, Header5, Header6
Data1.1, Data1.2,, Data1.4, Data1.5, Data1.6
Data2.1, Data2.2, Data2.3, Data2.4, Data2.5, Data2.6
Data3.1, Data3.2, Data3.3, Data3.4, Data3.5, Data3.6
Data4.1, Data4.2, Data4.3, Data4.4, Data4.5, Data4.6
```

Becomes a table named for the value in the -t flag

```
+-----+-----+-----+-----+-----+-----+
|Header1| Header2| Header3| Header4| Header5| Header6|
+-----+-----+-----+-----+-----+-----+
|Data1.1| Data1.2| "null"  | Data1.4| Data1.5| Data1.6|
|Data2.1| Data2.2| Data2.3| Data2.4| Data2.5| Data2.6|
|Data3.1| Data3.2| Data3.3| Data3.4| Data3.5| Data3.6|
|Data4.1| Data4.2| Data4.3| Data4.4| Data4.5| Data4.6|
```

Note: A string of "null" is actually placed in null data locations.

Example Configuration file

```

<?xml version="1.0" encoding="UTF-8"?>
<!--
    Copyright (c) 2010-2011 VIVO Harvester Team. For full list of contributors, please see the AUTHORS file
    provided.
    All rights reserved.
    This program and the accompanying materials are made available under the terms of the new BSD license which
    accompanies this distribution, and is available at http://www.opensource.org/licenses/bsd-license.html
-->
<Task>
    <!--INPUT -->
    <Param name="inputFile">person.csv</Param>

    <!--
    <% ===== PARAMETERS
    =====
    <% The source relational database
    information:
    <%
    <%
    <% <Param name="driver"> - A JDBC driver is a java class which handles the interface between the program and a
    given %>
    <% database.The driver needs to exist within the classpath, for the program to be able to use
    it. %>
    <%
    EXAMPLES:
    >
    <% <Param name="driver">org.h2.Driver<
    /Param> %>
    <%
    <%
    <% <Param name="connection"> - JDBC uses a connection string related to the Driver being
    used. %>
    <% It is in the general format "jdbc:somejdbcvender:other data
    needed" %>
    <% EXAMPLES
    :
    <% Microsoft SQL server : <Param name="connection">jdbc:jtds:sqlserver://127.0.0.1:8080/databasename<
    /Param> %>
    <% H2 database ( http://www.h2database.com ) : <Param name="connection">jdbc:h2:directory/location<
    /Param> %>
    <%
    <%
    <% <Param name="username"> - A valid login with proper
    permissions %>
    <% <Param name="password"> - The associated secure
    password %>
    <% These are what the fetch uses when accessing the given database. These must be valid otherwise the harvest
    run %>
    <% will not have access to the database. By default we use "sa" for system administrator and a blank
    password. %>
    <%
    <%
    -->
    <Param name="driver">org.h2.Driver</Param>
    <Param name="connection">jdbc:h2:data/csv/store</Param>
    <Param name="username">sa</Param>
    <Param name="password"></Param>
    <!-- TABLES -->

    <!--
    <% In this section the tables harvested from are
    described. %>
    <% <Param name="tableName"> This parameter the name of the table that is being harvested
    to. %>
    <%
    <%
    -->
    <Param name="tableName">CSV2</Param>
</Task>

```