



Name of Solution:

Informatica PowerCenter File Compare

Business Requirement:

Metadata-driven file comparison utility that automates regression testing of delimited flat files and significantly reduces the testing/comparison effort involved.

Solution URL:

<https://community.informatica.com/solutions/2682>

Supported Versions:

PowerCenter 8.6

Description:

Flat files are a popular mode of transferring data. At times there is a need to compare flat files generated in Data Warehousing / Data Integration environments (to determine the impact, if any, of a code change / enhancement / etc.), especially if these files happen to be reports shared with downstream users.

Informatica PowerCenter File Compare is a metadata-driven utility that helps automate regression testing of delimited flat files. Once the metadata of the files to be compared is made available to the utility, it can compare files of a similar layout and publish the summarized and detail test results. The results are categorized into three reports (tables)-

- Compare File Summary
- Compare Column Summary
- Compare Column Detail

The metadata to be provided includes the file layout, key columns and non-key columns to be compared / ignored.

Download file contents:

1. Workflow
2. Script
3. Metadata Query Generator

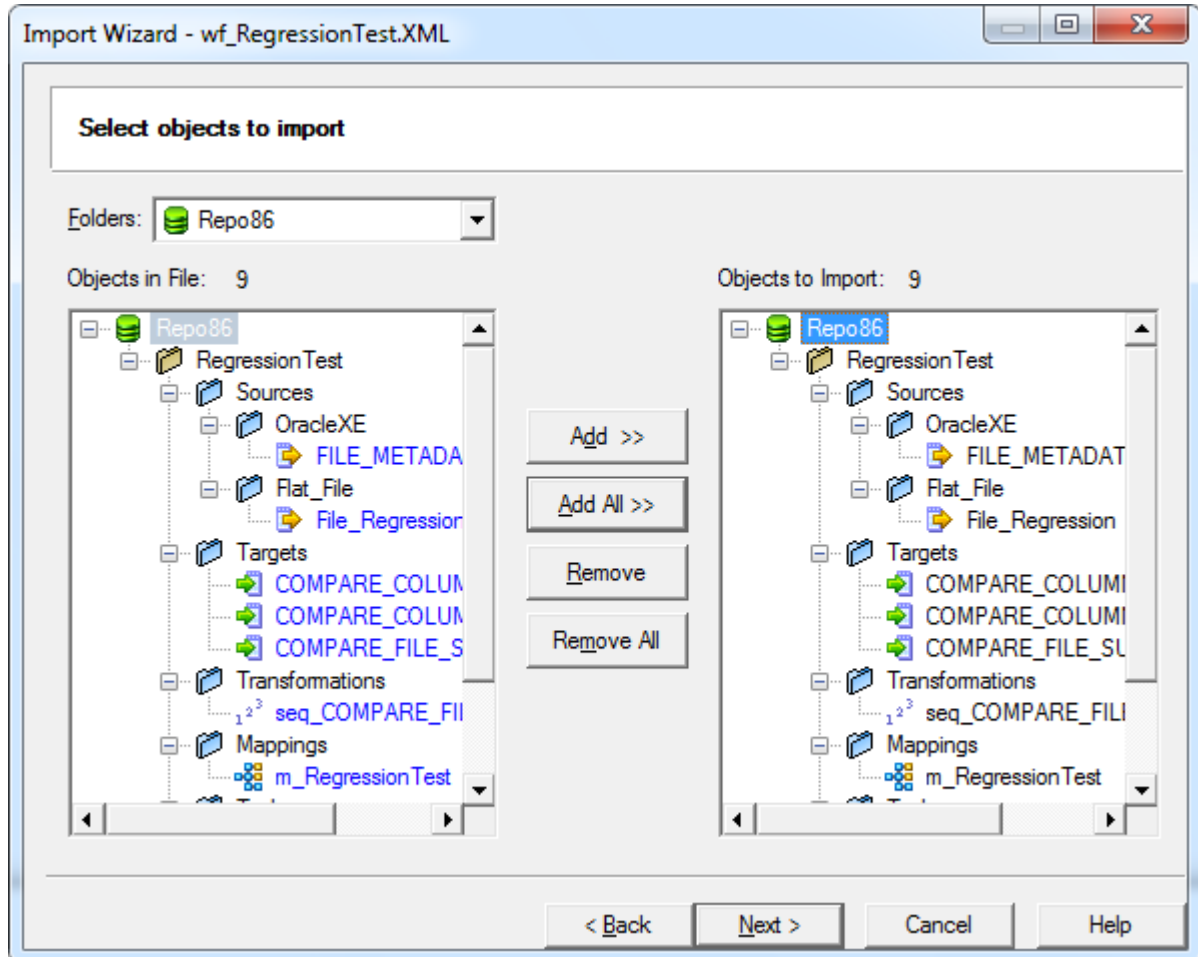
Steps to implement the solution:

Steps to implement the solution:

1. Create Tables
 - Using the DDLs provided, create the tables (FILE_METADATA, COMPARE_FILE_SUMMARY, COMPARE_COLUMN_SUMMARY and COMPARE_COLUMN_DETAIL) in the database.

informatica Marketplace

- If a non-Oracle database is being used, the syntax of the DDLs may have to be amended as they were formulated on Oracle.
2. Import Informatica Objects
- Select the appropriate folder in the repository
 - Import all the Informatica objects from the XML



For more details of importing object, please visit our [YouTube](#) link.

3. Build File Metadata
- Using the provided spreadsheet (QueryGenerator), create the file metadata insert queries; update only the green cells
 - File ID: Assign a unique integer value
 - File Name: Assign a unique name
 - Column Name: Specify the names in the order they appear in the file
 - Key Flag: Update as Y/N to indicate if the column is part of the key
 - Check Flag: Update as Y/N to indicate if non-key column is to compare/ignored.
 - Only non-key columns with Check Flag = 'Y' would be compared

Metadata Query Generator (Edit only the GREEN cells)						
File ID	2					
File Name	Employee					
File ID	File Name	Column ID	Column Name	Key Flag	Check Flag	Insert SQL
2	Employee	1	ID	Y	N	insert into file_metadata values(2,'Employee',1,'ID','Y','N');
2	Employee	2	NAME	N	Y	insert into file_metadata values(2,'Employee',2,'NAME','N','Y');
2	Employee	3	DEPT	N	Y	insert into file_metadata values(2,'Employee',3,'DEPT','N','Y');
2	Employee	4	DESIGNATION	N	Y	insert into file_metadata values(2,'Employee',4,'DESIGNATION','N','Y');
2	Employee	5	SUPERVISOR	N	N	insert into file_metadata values(2,'Employee',5,'SUPERVISOR','N','N');
2	Employee	6	DOB	N	Y	insert into file_metadata values(2,'Employee',6,'DOB','N','Y');

- Execute the queries against the database; Repeat this process for each set of files to be compared

Database Connection

OracleXE

Sticky Database

Schema

INFA

```
1 select * from file_metadata where file_id = 1
```

1:46 [46] INS

Auto Commit: ON

Encoding

Log

1: FILE_METADATA [5] x

	FILE_ID	FILE_NAME	COLUMN_ID	COLUMN_NAME	KEY_FLAG	CHECK_FLAG
1	1	Employee	1	ID	Y	N
2	1	Employee	2	NAME	N	Y
3	1	Employee	3	AGE	N	Y
4	1	Employee	4	GRADE	N	N
5	1	Employee	5	DEPT	N	Y

4. Amend Workflow

- Create a session for each set of files to be compared
- Choose the appropriate delimiter and specify the file location

5. Execute Workflow

- On execution, the specified files are compared and the test results are published into the tables COMPARE_FILE_SUMMARY, COMPARE_COLUMN_SUMMARY and COMPARE_COLUMN_DETAIL

6. Test Results

- COMPARE_FILE_SUMMARY: Holds file-level metrics. One row with a unique COMPARE_ID is created for each execution.

Database Connection

OracleXE

Schema

INFA

1select * from compare_file_summary where file_id = 1 and compare_id = 4059

1:75 [75] INS

Auto Commit: ON Encoding:

Log

1: COMPARE_FILE_SUMMARY [1] X

#	COMPARE_ID	COMPARE_DATE	FILE_ID	CNT_BASELINE	CNT_NEW	CNT_KEY_MATCH	CNT_NKEY_MATCH	CNT_NKEY_MISMATCH	CNT_ORPHAN_BASELINE	CNT_ORPHAN_NEW
1	4059	2014-01-19 22:29:04	1	11	10	9	5	4	2	

- COMPARE_COLUMN_SUMMARY: Holds column-level metrics. One row is created for each column in the file.

Database Connection: OracleXE | Sticky | Database: | Schema: INFA

```
1 select * from compare_column_summary where file_id = 1 and compare_id = 4059
```

1:29 [29] INS | Auto Commit: ON | Encoding: | Log | 1: COMPARE_COLUMN_SUMMARY [5] X

	COMPARE_ID	FILE_ID	COLUMN_ID	COLUMN_NAME	CNT_MATCH	CNT_MISMATCH
1	4059	1	1 ID		9	0
2	4059	1	2 NAME		7	2
3	4059	1	3 AGE		7	2
4	4059	1	4 GRADE		6	3
5	4059	1	5 DEPT		7	2

- COMPARE_COLUMN_DETAIL: Highlights the column value differences. One row is created for each column that differs on each record on the file.

Database Connection ☐ Sticky Database Schema

OracleXE INFA

1select * from compare_column_detail where file_id = 1 and compare_id = 4059

1:36 [36] INS

Log 1: COMPARE_COLUMN_DETAIL [6] x Auto Commit: ON Encoding

#	COMPARE_ID	FILE_ID	RECORD_KEY	COLUMN_ID	COLUMN_NAME	VALUE_BASELINE	VALUE_NEW
1	4059	1	1060	2 NAME	Brian	Brian Aquadro	
2	4059	1	1070	5 DEPT	IT	HR	
3	4059	1	1080	3 AGE	46	47	
4	4059	1	1090	2 NAME	Royala	Royala Furnis	
5	4059	1	1090	3 AGE	60	61	
6	4059	1	1090	5 DEPT	IT	HR	

YouTube Video on Importing and Configuring Workflows:

<http://www.youtube.com/playlist?list=PLRreK2jjigWBQ4NPfp0QWTxYDvInEqSJ>

Other Useful links:

[Mapping Bundles](#)

[Workflow Bundles](#)

[Informatica Tools Bundles](#)

[Informatica for Social Media Bundles](#)

[Debugging Tools Bundles](#)

[Visio Templates Bundles](#)

[B2B Templates and Projects Bundles](#)

[Data Quality Packs and Plans Bundles](#)

[Script Bundles](#)