Course : Talend Open Studio: Implementing Data Integration

Practical course - 3d - 21h00 - Ref. TOT
Price : 1840 € E.T.

★★★★☆ 4,2 / 5

BEST

Talend Open Studio (TOS) is a free, open-source ETL application designed for large-scale data synchronization between databases. This training course will teach you how to design, model, and develop TOS "jobs" to address a wide range of integration challenges. Please note: this course focuses exclusively on Talend Open Studio (the free, community, open-source version). It does not cover the commercial Talend Studio edition, which is subscription-based and integrated into the Qlik-Talend Cloud portal.

Teaching objectives

At the end of the training, the participant will be able to:

- Design and develop jobs in the application ETL Talend
- Optimize the jobs developed by using contexts and datasets.
- Carrying out more complex transformations by using variables, expressions, and joins
- Execute and debug a job, track execution statistics

Intended audience
Developers, project managers, decision-making IT consultants, database administrators.

Prerequisites
Good knowledge of RDBMs and SQL. Knowledge of the Java language is a plus.

Practical details
Hands-on work
Series of mini-projects used to design increasing elaborate Talend jobs.

Teaching methods
Theoretical contributions illustrated with example implementations, then consolidated by creating mini-Talend projects.

Course schedule

1. Presentation

- Data integration ETL solutions.
- TOS: Installation, user preferences. Reference documentation.
- Product philosophy. Designing a job.

Hands-on work
Getting started with the tool. Overview of the specifications of the generic project.

2. Model the need, design a first job

- Business Modeler. Job Designer.
- Main connections. CSV and XML components.
- Simple transformation components.
- Viewing the generated code, executing a job.

Exercise
Developing a job that sorts a CSV source, filtering data, and storing the result in an XML file.

3. Optimize job development

- Configuring reusable connections using metadata.
- Updating the metadata and propagating it to jobs, importing/exporting metadata.
- Configuring jobs with contexts.
- Externalizing context variables in ".properties" and ".ini" files.
- Creating and managing your own variables.
- Generating data sets for tests.

Exercise
Refactoring a job using metadata and contexts. Generating a test data set for this job.

4. Working with databases

- Supported databases and main components.
- Configuring operations on tables.
- Metadata and context of connecting to a database schema.
- Sharing connections and managing transactions.
- Creating queries using SQLBuilder.

Exercise
Reading and updating a data repository hosted on a MySQL server.

5. Working with multisource data

- Overview of the tMap component.
- Configuring input flows, creating joins.
- Carrying out more transformations using variables, expressions and joins.
- Qualifying data using filters.
- Generating multiple outputs, managing rejects.

Exercise
Consolidating multisource data and generating a warehouse.

6 Add-ons

- Breaking a job down into sub-jobs, using tRunJob. Launching jobs in the command line. Periodic execution.
- Debugging a job, tracking execution statistics
- tJasperOutput reporting.

Exercise
Generating a Jasper Report from a warehouse.

Dates and locations

REMOTE CLASS
2026 : 18 May, 18 May, 14 Oct., 14 Oct., 2 Dec., 2 Dec.

PARIS LA DÉFENSE
2026 : 18 May, 14 Oct., 2 Dec.