

Updated on 11/10/2024

Sign up

Talend DI (Data Integration) training

3 days (21 hours)

Presentation

Talend is the leading commercial provider of open source data integration software.

Talend is an open source data integration platform. It provides a variety of software and services for data integration, data management, enterprise application integration, data quality, cloud storage and Big Data.

In this course, you'll learn all about Talend, from fundamentals to advanced, including data integration and big data with examples.

Like all our training courses, this one will introduce you to the latest stable release and its new features: [Talend 8](#).

Objectives

- Mastering Talend in a Big Data environment
- Best practices for creating flexible, robust information systems
- Understanding the tool's philosophy
- Establish links between files, applications and databases
- Implementing Jobs
- Data mapping with Talend
- Creating connectable ETL tasks
- Filter, modify and unify data, fully control the ETL process
- Using Java to integrate data
- Understanding Talend's advanced concepts
- Read and write data to HDFS and NoSQL databases
- Transforming jobs with Pig and Hive
- Managing data quality
- Using Scoop to simplify relational database migration in Hadoop
- How to use the component library

Target audience

Anyone wishing to use Talend Studio to interact with Big Data systems.

Prerequisites

- Knowledge of Java and SQL
- Knowledge of databases

Talend Data Integration training program

INTRODUCTION

- Installation
- Context and metadata

DATA MAPPING

- Map data
- Data generation
- Output of tMap rejects
- Filtering mapped data
- Cartographic DB data

CONVERSION, EXTRACTION, ADHESION

- Regex expressions
- XML to CSV conversion
- CSV to XML conversion
- Advanced XML
- Membership data
- Data splitting
- Data unification

USEFUL FUNCTIONS

- Removing and adding properties
- Number of lines
- Temp, delete, move, rename

JAVA

- The basics of tJava

- tJavaRow component
- tJavaFlex shutter
- Customized uses

DATABASE

- Using Hyper SQL
- Analysis of table diagrams
- Schematic import
- Reading and writing data in a DB

DEBUGGING, OPERATION, CONSTRUCTION, PROGRAMMING

- Debugging
- Waste treatment
- Using tMsgbox
- Processing input rejects with tMap
- Creating jobs for Talend
- Adding a context from the command line
- Task planning

ADVANCED CONCEPTS

- Sub-jobs
- Customized components
- tFileExcel (customized components)
- tSystem
- tSCP

Further information

Companies concerned

This training course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.