

Opleiding : Talend Open Studio voor gegevenskwaliteit, gegevenskwaliteit beheren

Praktijkcursus - 2d - 14u00 - Ref. TDQ

Prijs : 1430 € V.B.

Open Studio for Data Quality, een open source tool voor dataprofilering. Je leert hoe je deze Talend tool effectief kunt gebruiken om het niveau van datakwaliteit in je informatiesysteem te beoordelen. Je implementeert analyses, meet of gegevens voldoen aan interne of externe standaarden en definieert strategieën voor het opschonen van foutieve gegevens. Deze training is uitsluitend voor Talend Open Studio (community, gratis en open source versie). Het omvat niet de commerciële versie van Talend Studio, onder betaalde licentie, geïntegreerd in het Qlik-Talend Cloud portaal.

Pedagogische doelstellingen

Aan het einde van de training is de deelnemer in staat om:

- ✓ Verbinding maken met gegevensbronnen, statistieken produceren, te profileren gegevens identificeren
- ✓ De verschillende soorten indicatoren en analyses kiezen die geschikt zijn voor de te controleren gegevens
- ✓ Complexe analyses uitvoeren om bedrijfsregels te verifiëren
- ✓ Strategieën definiëren voor het corrigeren van foutieve gegevens via Talend Data Integratiejobs

Doelgroep

Bedrijfsanalisten, data-integrators, datamanagers.

Voorafgaande vereisten

Goede kennis van relationele databases en SQL. Basiskennis van Talend Open Studio voor data-integratie.

Praktische modaliteiten

Leer methodes

70% van de tijd wordt besteed aan het gebruik van de tool. Elke deelnemer heeft zijn eigen werkplek.

Opleidingsprogramma

DEELNEMERS

Bedrijfsanalisten, data-integrators, datamanagers.

VOORAFGAANDE VEREISTEN

Goede kennis van relationele databases en SQL. Basiskennis van Talend Open Studio voor data-integratie.

VAARDIGHEDEN VAN DE CURSUSLEIDER

De deskundigen die de cursus leiden zijn specialisten op het betreffende vakgebied. Zij werden geselecteerd door onze pedagogische teams zowel om hun vakkennis als hun pedagogische vaardigheden voor elke cursus die zij geven. Zij hebben minstens vijf tot tien jaar ervaring in hun vakgebied en oefenen of oefenden verantwoordelijke bedrijfsfuncties uit.

BEOORDELINGSMODALITEITEN

De cursusleider beoordeelt de pedagogische vooruitgang van de deelnemer gedurende de gehele cursus aan de hand van meerkeuzevragen, praktijksituaties, praktische opdrachten, ...
De deelnemer legt ook van tevoren en naderhand een test af ter bevestiging van de verworven kennis.

1 Het probleem van gegevenskwaliteit

- De kwaliteit van gegevens in een informatiesysteem beoordelen.
- Fundamentele criteria: volledigheid, nauwkeurigheid en integriteit van gegevens.
- Talend Open Studio for Data Quality positioneren in de Talend suite.

Praktisch werk

Het product installeren, voorkeuren configureren.

2 De fundamentele concepten van TOS voor gegevenskwaliteit

- Metagegevens: verbindingen met databases, gedelimeerde bestanden en Excel-bestanden.
- Overzicht van de verschillende soorten analyses.
- Analysehulpmiddelen en -indicatoren.
- De gegevensverkenner.

Praktisch werk

Voer een eerste kolommenanalyse uit op gegevens uit een csv-bestand en analyseer de resultaten.

3 Eenvoudige analyses

- Zoeken naar duplicaten, naleving van intervalbeperkingen, datumnotatie, e-mailnotatie, enz.
- Tabel metriek, functionele afhankelijkheden tussen kolommen.
- Identificatie van overbodige waarden.
- Consistentiecontroles tussen foreign en primary keys.
- Gebruik indicatoren, modellen, regels en bronbestanden.

Praktisch werk

Voer een analyse van elk type uit op een set gedeeltelijk foutieve gegevens.

4 Geavanceerde analyses

- Analyse van schema en tabelstructuur met Data Explorer.
- Analyse van meerdere tabellen en kolommen, naleving van bedrijfsregels.
- Correlaties tussen kolommen zoeken en weergeven.
- Maak je eigen indicatoren en bronbestanden.
- Analyses beheren.

Praktisch werk

Maak een complexe bedrijfsregel die meerdere tabellen omvat en koppel hem aan een taak. Publiceer de regel in Talend Forge.

5 Geavanceerde elementen

- Gebruik contextvariabelen.
- Modellen maken op basis van reguliere expressies.
- Analyses exporteren/importeren en gegevens analyseren.
- Datafouten corrigeren met Talend Data Integratie.

Praktisch werk

Metadata en analyses instellen met behulp van contextvariabelen, geanalyseerde gegevens exporteren voor correctie in Talend Data Integration.

PEDAGOGISCHE EN TECHNISCHE MIDDELEN

- De gebruikte pedagogische middelen en cursusmethoden zijn voornamelijk: audiovisuele hulpmiddelen, documentatie en cursusmateriaal, praktische oefeningen en correcties van de oefeningen voor praktijkstages, casestudies of reële voorbeelden voor de seminars.
- Na afloop van de stages of seminars verstrekt ORSYS de deelnemers een evaluatievragenlijst over de cursus die vervolgens door onze pedagogische teams wordt geanalyseerd.
- Na afloop van de cursus wordt een presentielijst per halve dag verstrekt, evenals een verklaring van de afronding van de cursus indien de stagiair alle sessies heeft bijgewoond.

TOEGANGSMODALITEITEN EN TERMIJNEN

De inschrijving dient 24 uur voor aanvang van de cursus plaatsgevonden te hebben.

TOEGANKELIJKHEID VOOR MINDERVALIDEN

Is voor u speciale toegankelijkheid vereist? Neem contact op met mevr. FOSSE, contactpersoon voor mindervaliden, via het adres psh-accueil@ORSYS.fr om uw verzoek en de haalbaarheid daarvan zo goed mogelijk te bestuderen.

