

Opleiding : Apache Hop, gegevensstromen orkestreren

Je ETL-processen visueel beheren

Praktijkcursus - 3d - 21u00 - Ref. HOA

Prijs : 2010 € V.B.

NEW

Gegevens zijn van vitaal belang voor bedrijven. Apache Hop is krachtige, aanpasbare open source software voor het beheren van ETL-processen (Extract, Transform, Load). Het stelt je in staat om het verzamelen, transformeren en organiseren van gegevens uit heterogene bronnen te automatiseren en vervolgens naar een specifieke bron te sturen. Bespaar tijd door pijplijnen en workflows te beheren via een toegankelijke interface.

Pedagogische doelstellingen

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

- ✓ De Hop-omgeving en de werking ervan begrijpen en kunnen uitleggen
- ✓ Weten hoe je gegevens extraheert en transformeert
- ✓ De verschillende gegevensbronnen onderscheiden en organiseren
- ✓ Verschillende soorten gegevens combineren, classificeren en vergelijken
- ✓ Taakautomatisering toepassen en fouten analyseren om erop te reageren

Doelgroep

Iedereen die gegevensstromen voedt en verwerkt in een BI-database.

Voorafgaande vereisten

Goede kennis van SQL.

Praktische modaliteiten

Oefening

Toepassing van theorie op concrete gevallen, groepsdiscussies en praktische oefeningen.

Leer methodes

Actief lesgeven.

Opleidingsprogramma

DEELNEMERS

Iedereen die gegevensstromen voedt en verwerkt in een BI-database.

VOORAFGAANDE VEREISTEN

Goede kennis van SQL.

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 vak kennis 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 Apache Hop

- Waarom hop?
- Geschiedenis en presentatie in de breedste zin van het woord
- De tool installeren en configureren

Praktisch werk

Apache Hop installeren en configureren.

2 Een eerste gegevensextractie genereren

- Workflows en pijplijnen (planning)
- Van extractie tot gegevensinvoer
- Gegevensstromen begrijpen en beheren
- Een pijplijn en workflow uitvoeren

Praktisch werk

Een pijplijn ontwerpen en een workflow creëren.

3 Toegang tot bron- en doelgegevens

- Het concept van metadata
- Toegang tot gegevensbronnen configureren
- Ondersteunde bronnen/doelen
- Koppelingen tussen bronnen (joins)
- Voeding invoegen/bijwerken

Praktisch werk

Configureer toegang tot gegevensbronnen door ondersteunde doelen/bronnen en joins te identificeren.

4 Gegevens verwerken

- Sorteert je stroom omhoog of omlaag
- Je stroom verdubbelen
- Gegevens filteren op basis van verschillende criteria (de stroom verlichten)
- Informatie uit een veld halen (tekenreeks)
- Een gegevensitem vervangen door een ander
- Beheer van operators/berekeningen op de stroom
- Het Cartesisch product gebruiken
- Informatie uit heterogene gegevens koppelen
- Gegevensstromen vergelijken

Praktisch werk

Gegevensstromen verwerken door sorteren, splitsen en filteren van gegevens. Stromen vergelijken.

5 Uw gegevensstroom verrijken

- Logboek genereren
- Variabelen aanmaken en ophalen (datums, numeriek, alfanumeriek)
- Het resultaat van een stroom gebruiken
- De eigenschappen van een stroom en zijn planner

Praktisch werk

Maak variabelen aan en haal ze op, gebruik de resultaten van de flow en verbeter je gegevensstroom.

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.

6 De lussen

- De problemen
- Lussen met parameters en het "Kopieer regels naar resultaat" onderdeel
- Lussen met de "Kopieer rijen naar resultaat" en "Haal rijen op uit resultaat" componenten

Praktisch werk

De componenten van lussen begrijpen en manipuleren.

7 Werking

- Fouten beheren
- Logboeken genereren
- Fouten begrijpen en waarschuwingen starten (debuggen)
- Parallellisatie (gelijktijdige uitvoering van meerdere gegevensstromen)
- Ontwikkelingen import/export
- Taak/transformatie automatisering
- Documentatie (implementatie van standaarden in geval van fouten of rework)

Praktisch werk

Voer workflows gelijktijdig uit, beheer fouten en automatiseer taken.

Data en plaats

KLAS OP AFSTAND

2026 : 8 juni, 14 sep., 30 nov.

PARIS LA DÉFENSE

2026 : 1 juni, 7 sep., 23 nov.