

## Workshop Apache Spark

Bliksemsnel de beginselen van Apache Spark onder de knie krijgen? Daar is deze workshop voor. In deze 4 uur durende workshop leer je Apache Spark te gebruiken om Big Data te verwerken én om er Machine Learning (kunstmatige intelligentie) op toe te passen!

### Programma

We beginnen met een korte uitleg over wat Apache Spark precies is, hoe het werkt en wat je ermee kunt. Vervolgens wisselen we presentaties over de verschillende Apache Spark modules af met codeoefeningen. In deze oefeningen ga je zelf aan de slag met leuke opdrachten, zoals het verwerken van een livestream van Twitter-berichten.

#### Presentatie-onderdelen:

- Apache Spark (algemeen)
- Spark Core
- Spark SQL
- Spark Streaming
- Spark MLlib (machine learning)
- Spark GraphX

#### Zelf aan de slag:

- Zoeken in alle werken van Shakespeare
- Sentiment herkennen in Twitter-berichten
- Twitter-berichten verwerken uit een livestream
- Getallen in afbeeldingen herkennen (Machine Learning)

**Doelgroep:** De workshop is gericht op software engineers die nog nooit met Apache Spark gewerkt hebben. Tijdens de oefeningen wordt er in Java 8 geprogrammeerd. Alle voorbeelden van code in de presentatie zijn ook in Java geschreven.

## Voorvereisten

De deelnemers dienen ervaring te hebben met het programmeren in Java. Enige kennis van de Java 8 Stream API is nodig voor de oefeningen. De deelnemers dienen Docker te kunnen installeren en gebruiken op hun laptop.

Neem zelf de laptop mee waarop de te gebruiken software geïnstalleerd en geconfigureerd is. Ter voorbereiding op de workshop ontvang je uiteraard nog exacte instructies.

## Doelstellingen

### Na het volgen van deze workshop:

- Begrijp je wat Apache Spark is;
- Weet je wat je ermee kan doen;
- Heb je voldoende ervaring opgedaan in het werken met Apache Spark om er zelfstandig verder mee aan de slag te gaan.