



Scala and Big Data with Apache Spark

Introduction

As data grows at unprecedented rates, traditional tools often fail to keep pace. Apache Spark delivers the speed, scalability, and efficiency needed to handle big data processing and analysis at scale. With its distributed architecture and in-memory computation, Spark empowers developers to extract insights quickly and manage complex data pipelines with ease. Combined with the expressive power of Scala, Spark becomes an essential tool for anyone tackling large-scale data challenges. This seminar provides a jump start into Spark with Scala, giving participants the knowledge and confidence to begin working effectively in the big data era.

Prerequisites

This seminar was developed for software developers with extensive experience in the Scala programming language.

Duration

This seminar includes four academic hours. It can be delivered either in four academic hours or in two separate meetings of two academic hours each.

Premium Training

This seminar can be delivered either in Hebrew or in English. It can be delivered either online, or in a Hybrid way that allows the participants to choose whether to attend the meeting room where the training takes place or to join online. During the training itself we use a Moodle based Learning Management System you connect using a web browser from anywhere.

The Lecturer

This seminar is delivered by [Haim Michael](#), an experienced, well-known software development trainer with more than 25 years of experience in software development training, more than 30 years of experience in Java and more than 16 years of experience in Scala.

The Topics

This seminar sets the foundation for using Apache Spark. It covers the following topics (each topic has a separated dedicated presentation):

Introduction

Jump Start

Resilient Distributed Dataset

DataFrame API