

Tytuł szkolenia: Budowanie Data Pipelines

Kod szkolenia: Data-Pipelines

Wprowadzenie

Szkolenie z tworzenia data pipeliens w architekturze mikroserwisowej z Kafka, Cassandra i Spark.

Adresaci szkolenia

Cel szkolenia

- Nabycie umiejętności modelowania danych z bazą Cassandra;
- Poznanie języka z CQL i dobrych praktyk pracy z danymi;
- Zrozumienie natury i architektury systemów opartych o mikro usługi;
- Poznanie w praktyce zalet budowania usług w oparciu o klaster Kafka, bazę Cassandra i silnik Spark.

Czas i forma szkolenia

- 14 godzin (2 dni x 7 godzin), w tym wykłady i warsztaty praktyczne.

Plan szkolenia

1. Cassandra

- Charakterystyka bazy Cassandra
- Modelowanie danych, denormalizacja, klucze
- Praca z danymi w CQL

2. Mikroserwisy

- Architektura mikroserwisów
- Komunikacja między usługami, message broker

3. Kafka

- Charakterystyka i architektura klastra Kafka
- Praca z producers, consumers i streams

4. Kafka i Cassandra

- Budowanie usług z warstwą danych opartą Kafka i bazę Cassandra
- Praca z danymi strumieniowymi i bazą danych

5. Spark

- Charakterystyka i architektura silnika Spark
- Analiza danych z DataFrame oraz SQL API