

Zarządzanie danymi Big Data

Umiejętność gromadzenia i przetwarzania danych staje się coraz bardziej istotnym czynnikiem w rozwoju gospodarki i codziennym życiu. Big Data, czyli ogromne ilości zmieniających się danych jest odpowiedzią na to zapotrzebowanie. Gromadzenie i przetwarzanie wymaga elastycznego podejścia do dostarczania, przetwarzania i utrwalania informacji przy wykorzystaniu różnorodnych narzędzi. Przetwarzanie oraz analiza Big Data to z biznesowego punktu widzenia niezwykle wartościowe przedsięwzięcia.

W trakcie szkolenia pokazany zostanie sposób gromadzenia oraz przetwarzania różnych typów danych na specjalnie przygotowanym środowisku chmurowym. Szkolenie obejmuje wykłady, ćwiczenia oraz projekt, którego realizacja w sposób praktyczny uczy pracy ze zbiorami Big Data.



Przeznaczenie szkolenia

Szkolenie jest skierowane do architektów, projektantów rozwiązań, analityków i osób pracujących z danymi, którzy chcą poznać szeroki zakres możliwości, jakie niosą rozwiązania Big Data. Na szkoleniu podejmowana jest próba podziału rozwiązań Big Data na elementy składowe, poznania ich funkcjonalności oraz zestawienia tych elementów w całościowe rozwiązanie problemu. Omawiane są zarówno rozwiązania Open Source, jak również rozwiązania dostawców komercyjnych.



Korzyści wynikające z ukończenia szkolenia

- zrozumienie terminu Big Data

- nabycie umiejętności wykorzystywania różnych narzędzi do gromadzenia, przechowywania i przetwarzania danych
- zapoznanie się z zagadnieniem strumieniowego przetwarzania danych i obsługi zdarzeń
- wiedza na temat funkcjonowania komponentów wchodzących w skład platformy Hadoop
- analiza danych z wykorzystaniem ogólnie dostępnych narzędzi
- poznanie koncepcji (IoT) i porównanie z Big Data
- poznanie różnic pomiędzy bazami danych NoSQL a relacyjnymi bazami danych
- poznanie narzędzi dostarczania danych: Kafka, NiFi, MQTT



Oczekiwane przygotowanie słuchaczy

Na szkoleniu nie jest wymagana wiedza programistyczna. Ćwiczenia będą przygotowane w taki sposób, żeby osoba początkująca mogła je zrealizować. Niezrealizowanie jednego z zadań nie będzie wykluczać zakończenia z sukcesem całego projektu.



Język szkolenia

- Szkolenie: polski



Szkolenie obejmuje

FORMA SZKOLENIA

Szkolenie będzie zawierać wykład, ćwiczenia i projekt. Temat omówiony na wykładzie zostanie poruszony omawiany na ćwiczeniach oraz znajdzie się w realizowanym projekcie. Uczestnik we własnym zakresie będzie realizował projekt, budując jego poszczególne elementy w przygotowanym środowisku chmurowym.



Czas trwania

2 dni / 14 godzin

Agenda szkolenia

Dzień pierwszy

1. Wprowadzenie do rozwiązań chmurowych
2. Pozyskiwanie i transport danych (Wykład)
3. Wymiana danych (Ćwiczenia)
4. Przetwarzanie rozproszone MapReduce / HDFS (Wykład)
5. Systemy przetwarzania rozproszonego na przykładzie Hadoop-a (Ćwiczenia)
6. Strumieniowe przetwarzanie danych i zdarzeń (Wykład) / (Ćwiczenia)

Dzień drugi

1. Przechowywanie danych HDFS / Bazy danych NoSQL (Wykład)
2. Jak działa Hadoop na przykładzie jednej z dystrybucji (Ćwiczenia)
3. Bazy NoSQL (Ćwiczenia)
4. Przygotowanie infrastruktury chmurowej do ćwiczeń i projektu (Ćwiczenia)
5. Usługi chmurowe do przetwarzania danych (Ćwiczenia)
6. Analiza i prezentacja danych (Ćwiczenia)
7. Zakończenie projektu (Ćwiczenia)