

Szkolenie: Cloudera
Cloudera Search Training

| FORMA SZKOLENIA | MATERIAŁY SZKOLENIOWE | CENA | CZAS TRWANIA |
|------------------|-----------------------|-----------------|--------------|
| Stacjonarne | Cyfrowe | 1780 EUR NETTO* | 3 dni |
| Stacjonarne | Tablet CTAB | 1930 EUR NETTO* | 3 dni |
| Metoda dlearning | Cyfrowe | 1780 EUR NETTO* | 3 dni |
| Metoda dlearning | Tablet CTAB | 1780 EUR NETTO* | 3 dni |

* (+VAT zgodnie z obowiązującą stawką w dniu wystawienia faktury)

LOKALIZACJE

Kraków - ul. Tatarska 5, II piętro, godz. 9:00 - 16:00

Warszawa - ul. Bielska 17, godz. 9:00 - 16:00

Cel szkolenia:

Trzydniowe szkolenie **Cloudera Search** jest przeznaczone dla programistów i inżynierów danych, którzy chcą indeksować je na platformie **Hadoop** w celu wykonywania w czasie rzeczywistym ciężkich zapytań. Uczestnicy dowiedzą się, jak wydobyć więcej wartościowych informacji ze swoich danych poprzez integrację **Cloudera Search** z zewnętrznymi systemami.

Cloudera Search to pełnotekstowe, interaktywne wyszukiwanie oraz skalowalne, elastyczne indeksowanie na platformie **Hadoop** oraz rozwiązaniu enterprise data hub. Napędzany przez Apache Solr, Cloudera Search dostarcza skalowalność i niezawodność do najnowszej generacji zintegrowanych, wysokoobciążających zapytań.

Poprzez prowadzone przez instruktora dyskusje i interaktywne ćwiczenia praktyczne uczestnicy zdobędą wiedzę na temat ekosystemu aplikacji **Hadoop**, która zawiera:

- Wykonywanie wsadowego indeksowania danych zapisanych na HDFS i w Hbase.
- Wykonywanie strumieniowego indeksowania w czasie rzeczywistym przy pomocy Apache Flume.
- Indeksowanie treści w wielu językach i formatach plików.
- Procesowanie i transformowanie napływających danych przy pomocy Morphlines.
- Tworzenie interfejsów użytkownika dla indeksowanych danych przy użyciu Hue.
- Integrację Cloudera Search z zewnętrznymi aplikacjami.
- Udoskonalanie interakcji z Cloudera Search używając funkcjonalności, takich jak: wyłuszczenie, wyszukiwanie przy użyciu dyskretnych kryteriów i uwzględniania nieprawidłowej pisowni.

Plan szkolenia:

- Wprowadzenie
- Wstęp do Cloudera Search
 - Co to jest Cloudera Search?
 - Pomocne funkcjonalności
 - Przypadki użycia
 - Podstawowa architektura
- Wykonywanie prostych zapytań
 - Wykonywanie zapytań z poziomu Admin UI
 - Podstawowa składnia
 - Techniki przybliżonych dopasowań
 - Kontrolowanie rezultatów zapytań
- Tworzenie bardziej skomplikowanych zapytań
 - Miara znaczenia i filtry
 - Parsery zapytań
 - Funkcje
 - Wyszukiwanie geoprzestrzenne
 - Dyskretne kryteria wyszukiwania
- Przygotowanie do indeksowania dokumentów
 - Wstęp do procesu indeksowania
 - Jak używać Morphlines?
 - Tworzenie plików konfiguracyjnych
 - Projektowanie schematów
 - Zarządzanie kolekcjami
- Indeksowanie wsadowe danych na HDFS Data przy użyciu MapReduce
 - Wstęp do wsadowego indeksowania na HDFS
 - Używanie MapReduce Indexing Tool
 - Testowanie i rozwiązywanie problemów
- Indeksowanie w czasie prawie-rzeczywistym używając Flume
 - Wstęp do indeksowanie w czasie prawie-rzeczywistym
 - Wprowadzenie do Apache Flume
 - Jak wykonać Indeksowanie w czasie prawie-rzeczywistym, wykorzystując Flume?
 - Testowanie i rozwiązywanie problemów
- Indeksowanie danych w HBase Data przy pomocy Lily
 - Co to jest Apache HBase?

- Indeksowanie wsadowe dla HBase
- Indeksowanie danych w HBase Tables w czasie prawie-rzeczywistym
- Indeksowanie danych w różnych językach i formatach
 - Typy pól i sekwencje analizatorów
 - Stemming słów, mapowanie znaków oraz wsparcie dla różnych języków
 - Wsparcie dla schematów i analiz w Admin UI
 - Wydobywane treści i metadanych używając Apache Tika
 - Indeksowanie plików binarnych w SolrCell
- Usprawnienie jakości i wydajności wyszukiwania
 - Dostarczanie dopasowanych wyników
 - Pomaganie użytkownikom w zdobyciu informacji
 - Wydajność i rozwiązywanie problemów z zapytaniami
- Tworzenie interfejsów użytkownika dla Cloudera Search
 - Wstęp do Search UI
 - Budowanie interfejsów użytkownika używając Hue
 - Integracja Cloudera Search we własnych aplikacjach
- Rozważania na temat wdrażania
 - Planowanie wdrożenia
 - Ustalanie wymagań sprzętowych
 - Wstęp do bezpieczeństwa
 - Aliasowanie kolekcji
- Podsumowanie

Wymagania:

Kurs jest przygotowany dla programistów i inżynierów danych z co najmniej podstawową znajomością technologii Hadoop i doświadczeniem w programowaniu w standardowych językach programowania tj. [Java](#), C, C++, [Perl](#) lub [Python](#). Uczestnicy powinni sprawnie poruszać się po linii poleceń systemu Linux i potrafić wykonać podstawowe zadania, jak tworzenie i usuwanie katalogów, wyświetlanie i zmiana uprawnień do plików, uruchamianie skryptów i analiza rezultatu ich wykonania. Nie jest konieczne wcześniejsze doświadczenie z Apache Solr lub Cloudera Search ani z HBase lub SQL.

Poziom trudności



Certyfikaty:

Uczestnicy otrzymają **certyfikaty** uczestnictwa sygnowane przez **Cloudera**.

Prowadzący:

Certyfikowany instruktor Cloudera