

# POTOK BIG DATA PRZYPADKI UŻYCIA

## Contents

Wstęp	1
Budowanie kompetencji	1
Wymagania	1
Cele kursu	2
Potok Big Data	2
Zastosowania Big Data	4
Bibliografia	4

## 1. Wstęp

Ten wykład przedstawia potok Big Data i jego komponenty oraz przedstawia ich główne cechy i zastosowanie aplikacji.

## 2. Budowanie kompetencji

Rozważenie i opanowanie tego tematu zbuduje umiejętności i wiedzę uczniów związaną z:

- Architekturą systemu Big Data
- Komponentami Big Data
- Etapami potoku Big Data
- Technologie Big Data do wdrożenia potoku

## 3. Wymagania

**Disclaimer:** Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency (NA). Neither the European Union nor NA can be held responsible for them.



Wiedza podstawowa studentów do przedmiotu wykładu:

Podstawy systemów operacyjnych, struktur danych, algorytmów, technologii przechowywania i zarządzania danymi.

#### 4. Cele kursu

Po ukończeniu tego wykładu uczniowie:

- Zapoznają się z koncepcjami i architekturami potoku Big Data
- Zrozumieją zasady, cechy i funkcje etapów i komponentów potoku Big Data, a także technologii ich implementacji
- Wykorzystają różne technologie w tworzeniu potoków Big Data

#### 5. Potok Big Data

##### **Potok Big Data: technologie i techniki**

Narzędzia do zarządzania przepływem pracy, takie jak **Airflow** i **Luigi**, porządkują procesy tworzące potok, automatycznie rozwiązują zależności i dają programistom możliwość wizualizacji i organizowania przepływów pracy danych.

Struktury zdarzeń i komunikatów, takie jak **Apache Kafka** i **RabbitMQ**, przechwytyują zdarzenia z aplikacji biznesowych, udostępniając je jako strumienie o wysokiej przepustowości i umożliwiają komunikację między różnymi systemami przy użyciu własnych protokołów.

Terminowe planowanie procesów jest krytyczne w każdym potoku danych.

Wiele narzędzi umożliwia użytkownikom tworzenie szczegółowych harmonogramów dotyczących pozyskiwania, przekształcania i ładowania danych do miejsc docelowych, od prostego narzędzia cron po całe dedykowane platformy automatyzacji obciążeń.

##### **Przypadki użycia**

Zarządzanie: Przewidywanie ekonomicznych następstw niektórych wydarzeń, np. upadłości konkretnej firmy

Teledetekcja: Automatyczne wykrywanie zdarzeń rolniczych na podstawie obrazów satelitarnych

IoT i inteligentne środowiska

Monitorowanie transportu i ruchu

Spersonalizowana opieka zdrowotna: polecenie spersonalizowanych zabiegów, dawek leków w oparciu o profil genów

Zarządzanie z prognozowaniem i Przemysł 4.0: Przewiduj awarie i anomalie urządzeń

Statystyki zorientowane na dane: oszacuj wolne mieszkania na podstawie zużycia energii elektrycznej w czasie rzeczywistym

### **Potoki Big Data według branży**

Sklepy internetowe i stacjonarne, które śledzą trendy konsumenckie.

Organizacje opieki zdrowotnej, które analizują ogromne ilości danych w celu znalezienia skutecznych metod leczenia.

Instytucje bankowe i finansowe, które wykorzystują big data do przewidywania trendów i poprawy obsługi klienta.

Firmy budowlane, które śledzą wszystko, od kosztów materiałów po przepracowane godziny.

Firmy transportowe, które analizują ruch i pomagają osobom dojeżdżającym do pracy dotrzeć do celu tak szybko, jak to możliwe.

Organizacje zajmujące się rozrywką, mediami i komunikacją wykorzystują duże zbiory danych na kilka sposobów, takich jak udostępnianie mediów społecznościowych w czasie rzeczywistym i ulepszanie strumieniowego przesyłania multimediiów w jakości HD.

### **Potoki Big Data według branży (kontynuacja)**

Szkoły, kolegia i uniwersytety mierzą demografię studentów, przewidują trendy naboru, poprawiają sukcesy uczniów i określają, którzy nauczyciele się wyróżniają.

Grupy produkcyjne i zajmujące się zasobami naturalnymi potrzebują potoków danych big data, aby usprawnić swoje działania w celu obniżenia kosztów ogólnych, dostarczać produkty, których potrzebują konsumenci, i identyfikować potencjalne zagrożenia.

Rząd wykorzystuje potoki danych big data na wiele różnych sposobów, takich jak analiza danych w celu śledzenia zmian w środowisku, wykrywania oszustw, rozpatrywania wniosków o niepełnosprawność i identyfikowania chorób, zanim dotkną tysiące ludzi.

Firmy energetyczne wykorzystują potoki danych big data do zarządzania pracownikami w sytuacjach kryzysowych, szybkiego identyfikowania problemów, aby mogły zacząć szukać rozwiązań i dostarczać konsumentom niezbędnych informacji.

## **6. Zastosowania Big Data**

Analiza klientów w branży sprzedażowej

- Np. w celu zwiększenia retencji i lojalności klientów

Zarządzanie z prognozowaniem dla Przemysłu 4.0

- Np. wykrywanie nietypowych stanów maszyny w celu zmniejszenia kosztów utrzymania

Zapobieganie przestępczości

- Analiza wzorców i trendów przestępczości

Opieka zdrowotna

- Np. do diagnozowania i leczenia chorób genetycznych

Finanse

- Przewidywanie zachowań klientów i tworzenie strategii dla banków i instytucji finansowych

### **Kierunki zastosowania Big Data**

Sektor rządowy, m.in. przy użyciu otwartych danych

- <https://www.europeandataportal.eu/>

Edukacja

- Np. w celu usprawnienia procesu uczenia się, zaprojektowania nowego kursu

Astronautyka

- Np. odkrycia astronomiczne

**Big Data w opiece zdrowotnej**

**Big Data do analizy mediów społecznościowych**

**Potok Big Data w monitoringu transportu**

**Big Data w handlu**

## **7. Bibliografia**

- Nathan Marz, James Warren, Big Data Principles and best practices of scalable realtime data systems. Manning Publications Co., 2015.
- Martin Kleppmann, Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems, O'Reilly Media, 2017.

- Dean Wampler, *Fast Data Architectures for Streaming Applications*, 2nd edition, O'Reilly Media, 2019.
- A. Bahga, V. Madisetti, *Big Data Science and Analytics: A Hands-On Approach*, VPT, 2016.