

Liczba godzin zajęć

- Wykład: 15
- Ćwiczenia: 30

Wymagania wstępne:

Przedmiot wymaga:

1. podstawowa wiedza i umiejętności w zakresie programowania (najlepiej w języku Python, gdyż w tym języku prowadzone będą zajęcia).
2. podstawowa wiedza i umiejętności w zakresie relacyjnych baz danych.

Efekty kształcenia:

Po zakończeniu kursu student:

1. zna pojęcie dużego zbioru danych;
2. zna specyfikę przetwarzania dużych zbiorów danych;
3. potrafi pozyskiwać informacje z różnych źródeł na temat dużych zbiorów danych, integrować je, dokonywać ich interpretacji, a także wyciągać wnioski oraz w sposób jasny i czytelny formułować i uzasadniać opinie;
4. umie stosować poznane technologie w celu przetwarzania dużych zbiorów danych;
5. potrafi wykorzystywać poznane narzędzia i techniki do pozyskiwania, przechowywania i przetwarzania danych;
6. ma świadomość, iż duże zbiory danych dają ogromne możliwości, ale i pociągają za sobą ogromną odpowiedzialność ze szczególnym uwzględnieniem ochrony danych i skutków ekonomicznych i społecznych za podejmowane na ich podstawie decyzje.

Weryfikacja efektów: 1,6: egzamin; 2, 4, 5, 6: ćwiczenia; 3: projekt/opracowanie.

Skrócony opis:

Przedmiot zakłada podstawową wiedzę i umiejętności w zakresie baz danych i programowania. Celem wykładu jest przedstawienie ogólnego spojrzenia na zagadnienia

związane z przetwarzaniem dużych zbiorów danych. Celem ćwiczeń jest możliwie szerokie przedstawienie problematyki i specyfiki przetwarzania dużych zbiorów danych, z odniesieniami do typowych technologii jakie są w danym czasie wykorzystywane.

Sposoby i kryteria oceniania:

- Egzamin: **NIE**
- Wykład: zaliczenie (wymagana ocena pozytywna)

Ocena z wykładu jest oceną uzyskaną z testu zamkniętego oraz ewentualnych pytań problemowych. Odpowiedzi do części testowej są typu prawda/fałsz z punktami ujemnymi za błędną odpowiedź (+1 za odpowiedź poprawną, 0 za brak odpowiedzi, -1 za odpowiedź błędną). Z tej części można otrzymać od -N do N punktów. Punktacja pytań problemowych zależy od stopnia ich trudności (w tym przypadku będzie można uzyskać tylko punkty nieujemne). Z tej części można otrzymać od 0 do M punktów. Łącznie z egzaminu można otrzymać od -N+M do N+M punktów.

Kryteria oceniania zaliczenia wykładu::

- od 86% z N+M 5
 - od 72% z N+M 4+
 - od 58% z N+M 4
 - od 44% z N+M 3+
 - od 30% z N+M 3
 - poniżej 30% z N+M 2
- Ćwiczenia:
 - na ocenę 3.0-3.5 składa się zaliczenie poszczególnych ćwiczeń praktycznych oraz pozytywna ocena z projektu śródsemestralnego;
 - na ocenę 4.0-5.0 wymagane jest wykonanie projektu końcowego.
 - Przedmiot: średnia z wykładu i ćwiczeń

Metody dydaktyczne:

Przedmiot jest realizowany w postaci wykładu i ćwiczeń.

Treści kształcenia:

Przewidywane treści kształcenia.

Wykład

- **Zagadnienie 1** Big Data - koncepcje i terminologia.
- **Zagadnienie 2** Paradygmaty Big Data.
- **Zagadnienie 3** Koncepcje przetwarzania danych typu Big Data.
- **Zagadnienie 4** Big Data - modele danych.
- **Zagadnienie 5** Kolejki komunikatów.
- **Zagadnienie 6** Czy faktycznie Big Data jest nam potrzebne?
- **Zagadnienie 7** Problematyka adopcji Big Data.

Ćwiczenia

- **Zagadnienie 1.** Czyszczenie danych.
- **Zagadnienie 2.** Przetwarzanie dokumentów semistrukturalnych przy pomocy wyrażeń regularnych.
- **Zagadnienie 3.** Przetwarzanie wielowątkowe / wielozadaniowe - sformułowanie problemu.
- **Zagadnienie 4.** Kolejki komunikatów i zadań.
- **Zagadnienie 5.** Omówienie projektów zaliczeniowych.
- **Zagadnienie 6.** Podstawy pracy z nierelacyjną bazą danych.

Literatura:

1. Fulmański P., Engineering of big data processing, 2025.
2. Erl T., Khattak W., Buhler P., Big Data Fundamentals: Concepts, Drivers & Techniques, Part of the The Prentice Hall Service Technology Series from Thomas Erl series. Published Jan 5, 2016 by Prentice Hall.
3. Marz N. and Warren J., Big Data. Principles and best practices of scalable realtime data systems. Manning Publications, April 2015.

Punkty ECTS

Liczba godzin pracy studenta związanych z osiągnięciem efektów uczenia się w ramach zajęć zgodnie z planem studiów	Obciążenie pracą student				
	W	Ćw	L	Zk	P
zapoznanie z literaturą	15	0	0	30	0
przygotowanie do zajęć	5				
przygotowanie do kolokwium				5	
realizacja zadanych ćwiczeń i zadań					
przygotowanie sprawozdań z zajęć					
przygotowanie projektu/pracy zaliczeniowej				20	
przygotowanie do egzaminu					
Razem	20	0	0	55	0