

# Big data technologies - opis przedmiotu

| Informacje ogólne   |                                                                   |
|---------------------|-------------------------------------------------------------------|
| Nazwa przedmiotu    | Big data technologies                                             |
| Kod przedmiotu      | 11.3-WE-BizEIP-TechBigData-Er                                     |
| Wydział             | <a href="#">Wydział Informatyki, Elektrotechniki i Automatyki</a> |
| Kierunek            | Biznes elektroniczny                                              |
| Profil              | praktyczny                                                        |
| Rodzaj studiów      | Program Erasmus pierwszego stopnia                                |
| Semestr rozpoczęcia | semestr zimowy 2022/2023                                          |

| Informacje o przedmiocie        |                                                                                       |
|---------------------------------|---------------------------------------------------------------------------------------|
| Semestr                         | 3                                                                                     |
| Liczba punktów ECTS do zdobycia | 5                                                                                     |
| Typ przedmiotu                  | obowiązkowy                                                                           |
| Język nauczania                 | angielski                                                                             |
| Sylabus opracował               | <ul style="list-style-type: none"><li>dr hab. inż. Artur Gramacki, prof. UZ</li></ul> |

| Formy zajęć |                                         |                                        |                                            |                                           |                     |
|-------------|-----------------------------------------|----------------------------------------|--------------------------------------------|-------------------------------------------|---------------------|
| Forma zajęć | Liczba godzin w semestrze (stacjonarne) | Liczba godzin w tygodniu (stacjonarne) | Liczba godzin w semestrze (niestacjonarne) | Liczba godzin w tygodniu (niestacjonarne) | Forma zaliczenia    |
| Wykład      | 30                                      | 2                                      | -                                          | -                                         | Zaliczenie na ocenę |
| Projekt     | 30                                      | 2                                      | -                                          | -                                         | Zaliczenie na ocenę |

## Cel przedmiotu

Teaching students how to choose the right data analysis techniques depending on the scale of the problem being considered and the type of analysis being carried out.

Teaching students to work using modern platforms for data storage and processing.

Teaching students selected techniques to analyze large data sets, mainly textual.

## Wymagania wstępne

Introduction to databases, Basics of statistics

## Zakres tematyczny

**Big Data:** An introduction to processing large amounts of data.

**Non-relational databases:** Reminder of the basic issues related to relational databases. Advantages and disadvantages of these databases. Basic problems related to the use of relational databases to store and process larger and larger amounts of increasingly distributed data. Horizontal and vertical scaling of databases. A new concept of databases not based on the traditional relational model. CAP and BASE theory. Aggregate data models. Key-value, column, document and graph databases. Database replication. Sharing resources in databases. Map-Reduce methodology. Presentation of a few selected non-relational database systems (e.g. MongoDB, Cassandra, Redis, Neo4J, Oracle NoSQL Database).

**Selected IT systems:** Large-scale business analytics: modern solutions used for transmission, storage and processing of large data sets. Basics of data processing using convolutional neural networks (CNN). Tensorflow and Keras libraries. Working in the Google Colaboratory cloud environment.

## Metody kształcenia

lecture: conventional lecture

project: work in groups, practical classes

## Efekty uczenia się i metody weryfikacji osiągania efektów uczenia się

| Opis efektu                                                                                                                                      | Symbole efektów | Metody weryfikacji                                                                                                            | Forma zajęć                                             |
|--------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|-------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------|
| Knows the databases known as NoSQL, is able to define their basic features and types, is able to use the most popular systems of such databases. |                 | <ul style="list-style-type: none"><li>bieżąca kontrola na zajęciach</li><li>test egzaminacyjny z progami punktowymi</li></ul> | <ul style="list-style-type: none"><li>Wykład</li></ul>  |
| Is able to define the concept of Text Mining and give typical examples of tasks in this area of knowledge.                                       |                 | <ul style="list-style-type: none"><li>test egzaminacyjny z progami punktowymi</li></ul>                                       | <ul style="list-style-type: none"><li>Wykład</li></ul>  |
| Knows the techniques of searching for text information and creating their ranking.                                                               |                 | <ul style="list-style-type: none"><li>test egzaminacyjny z progami punktowymi</li></ul>                                       | <ul style="list-style-type: none"><li>Wykład</li></ul>  |
| Is able to perform selected statistical analyzes of text documents.                                                                              |                 | <ul style="list-style-type: none"><li>bieżąca kontrola na zajęciach</li></ul>                                                 | <ul style="list-style-type: none"><li>Projekt</li></ul> |

| Opis efektu                                                                                                                 | Symbole efektów Metody weryfikacji                                                          | Forma zajęć                                                 |
|-----------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|-------------------------------------------------------------|
| Is able to characterize the features of modern Big Data platforms.                                                          | <ul style="list-style-type: none"> <li>• test egzaminacyjny z progami punktowymi</li> </ul> | <ul style="list-style-type: none"> <li>• Wykład</li> </ul>  |
| Is able to use the tools offered by the Elasticsearch platform to perform advanced data analysis and real-time exploration. | <ul style="list-style-type: none"> <li>• bieżąca kontrola na zajęciach</li> </ul>           | <ul style="list-style-type: none"> <li>• Projekt</li> </ul> |

## Warunki zaliczenia

## Literatura podstawowa

1. Pramod J. Sadalage and Martin Fowler: NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence, 2012
2. Dan Sullivan: NoSQL for Mere Mortals, 2015
3. Francois Chollet: Deep Learning with Python, Helion, 2017
4. Tensorflow and Keras docs: <https://www.tensorflow.org/guide>

## Literatura uzupełniająca

## Uwagi

Zmodyfikowane przez dr hab. inż. Marek Kowal, prof. UZ (ostatnia modyfikacja: 06-04-2022 09:00)

Wygenerowano automatycznie z systemu SylabUZ