*Załącznik nr 4 do Zarządzenia Nr…………..*

# **KARTA KURSU**

|  |  |
| --- | --- |
| Nazwa | Data mining i korpusy - Analiza danych językowych |
| Nazwa w j. ang. | Data-mining and Corpora |

|  |  |  |
| --- | --- | --- |
| Koordynator | Dr Joanna Podhorodecka | Zespół dydaktyczny |
| Dr Joanna Podhorodecka  Dr Anna Ścibior-Gajewska |
|  |  |
| Punktacja ECTS\* | 4 |

Opis kursu (cele uczenia się)

|  |
| --- |
| Celem kursu jest:  - zdobycie przez uczestników umiejętności analizy dużych próbek językowych oraz umiejętności pracy na korpusie  - Zdobycie umiejętności przygotowania danych i przeprowadzenia analizy korelacyjnej w środowisku statystycznym R  Kurs prowadzony jest w języku angielskim. |

Warunki wstępne

|  |  |
| --- | --- |
| Wiedza | Podstawy wiedzy i umiejętność zastosowania podstawowej terminologii z zakresu językoznawstwa. |
| Umiejętności | Rozumienie tekstu pisanego i mówionego w języku angielskim na poziomie B2 |
| Kursy | -- |

Efekty uczenia się

|  |  |  |
| --- | --- | --- |
| Wiedza | Efekt uczenia się dla kursu | Odniesienie do efektów kierunkowych |
| W01 ma uporządkowaną pogłębioną wiedzę obejmującą terminologię, teorie i metodologię z zakresu filologii.  W02 zna i rozumie zaawansowane metody analizy, interpretacji, wartościowania i problematyzowania różnego rodzaju danych językowych właściwe dla wybranych tradycji, teorii lub szkół badawczych w zakresie filologii. | K2\_W01; K2\_W02  K2\_W05; K2\_W07 |

|  |  |  |
| --- | --- | --- |
| Umiejętności | Efekt uczenia się dla kursu | Odniesienie do efektów kierunkowych |
| U01 potrafi wyszukiwać, analizować, oceniać, selekcjonować i użytkować dane językowe z wykorzystaniem różnych źródeł oraz formułować na tej podstawie krytyczne sądy.  U02 potrafi przeprowadzić krytyczną analizę i interpretację danych językowych,, stosując oryginalne podejścia, uwzględniające nowe osiągnięcia humanistyki, w celu określenia ich znaczeń, oddziaływania społecznego, miejsca w procesie historyczno-kulturowym. | K2\_U01; K2\_U05  K2\_U02; K2\_U011 |

|  |  |  |
| --- | --- | --- |
| Kompetencje społeczne | Efekt uczenia się dla kursu | Odniesienie do efektów kierunkowych |
| K01 rozumie potrzebę krytycznej analizy danych językowych. | K2\_K01; K2\_K02 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Organizacja | | | | | | | | | | | | | |
| Forma zajęć | Wykład  (W) | Ćwiczenia w grupach | | | | | | | | | | | |
| A |  | K |  | L |  | S |  | P |  | E |  |
| Liczba godzin |  |  | | 20 | |  | |  | |  | | 10 | |

Opis metod prowadzenia zajęć

|  |
| --- |
| Wykład i prezentacja prowadzącej; dyskusja zadanych tekstów; praca projektowa w grupach i indywidualnie; Platforma Moodle.  W okresie nauczania zdalnego w czasie epidemii kurs odbywa się na platformach Teams (zajęcia synchroniczne) oraz Moodle (realizacja kolejnych etapów projektu przez funkcję Forum). |

Formy sprawdzania efektów uczenia się

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | E – learning | Gry dydaktyczne | Ćwiczenia w szkole | Zajęcia terenowe | Praca laboratoryjna | Projekt indywidualny | Projekt grupowy | Udział w dyskusji | Referat | Praca pisemna (esej) | Egzamin ustny | Egzamin pisemny | Inne |
| W01 | X |  |  |  |  | x | x | x |  | x |  |  |  |
| W02 | X |  |  |  |  | x | x |  |  |  |  | X |  |
| U01 | X |  |  |  |  | x | x |  |  | x |  | X |  |
| U02 | X |  |  |  |  | x | x |  |  | x |  | X |  |
| K01 | X |  |  |  |  | x | x |  |  |  |  |  |  |

|  |  |
| --- | --- |
| Kryteria oceny | Studenci oceniani są na podstawie aktywności na zajęciach oraz przygotowanych projektów:  ZALICZENIE na podstawie złożenia raportu z małego projektu (email)  EGZAMIN w formie projektu z raportem (Moodle). Standardowa skala ocen.  W okresie nauczania zdalnego w czasie epidemii wymagane jest regularne uczestnictwo w zajęciach online i systematyczne zamieszczanie kolejnych etapów realizacji projektu na forum Moodle. |

|  |  |
| --- | --- |
| Uwagi |  |

Treści merytoryczne (wykaz tematów)

|  |
| --- |
| Data mining: definicja, założenia, zastosowanie, cele;  Natural language processing; indexing, retrieval, information structure;  Large language samples analysis: stylometry; corpus research; internet as a language corpus;  Language corpora: Coca, BNC, NKJP (Polcra), PWN SJP;  COCA – research: project design;  Multiple Correspondence Analysis; R-statistical environment;  Project reports – discussion |

Wykaz literatury podstawowej

|  |
| --- |
| Reference Guide for the British National Corpus <http://www.natcorp.ox.ac.uk/docs/URG/>  The Corpus of Contemporary American English (COCA) and the British National Corpus (BNC) http://corpus.byu.edu/coca/compare-bnc.asp |

Wykaz literatury uzupełniającej

|  |
| --- |
| Jackson P & Moulinier I. 2002 *Natural Language Processing for Online Applications*, John Benjamins  Chakrabarti S. *Mining the Web*, Morgan Kaufman |

Bilans godzinowy zgodny z CNPS (Całkowity Nakład Pracy Studenta)

|  |  |  |
| --- | --- | --- |
| liczba godzin w kontakcie z prowadzącymi | Wykład |  |
| Konwersatorium (ćwiczenia, laboratorium itd.) | 30 |
| Pozostałe godziny kontaktu studenta z prowadzącym | 10 |
| liczba godzin pracy studenta bez kontaktu z prowadzącymi | Lektura w ramach przygotowania do zajęć | 10 |
| Przygotowanie krótkiej pracy pisemnej lub referatu po zapoznaniu się z niezbędną literaturą przedmiotu | 20 |
| Przygotowanie projektu lub prezentacji na podany temat (praca w grupie) | 30 |
| Przygotowanie do egzaminu/zaliczenia | 20 |
| Ogółem bilans czasu pracy | | 120 |
| Liczba punktów ECTS w zależności od przyjętego przelicznika | | 4 |