

WYDZIAŁ Informatyki i Zarządzania/ STUDIUM.....

**KARTA PRZEDMIOTU**Nazwa w języku polskim *Inżynieria języka naturalnego*Nazwa w języku angielskim *Natural Language Engineering*Kierunek studiów (jeśli dotyczy): *Informatyka*Specjalność (jeśli dotyczy): *Inteligentne Systemy Informatyczne*Stopień studiów i forma: **II stopień, stacjonarna**Rodzaj przedmiotu: **obowiązkowy**Kod przedmiotu **INZ003782**Grupa kursów **NIE**

|   | Wykład               | Ćwiczenia                      | Laboratorium                   | Projekt              | Seminarium                     |
|---|----------------------|--------------------------------|--------------------------------|----------------------|--------------------------------|
| Liczba godzin zajęć zorganizowanych w Uczelni (ZZU)   | 30                   |                                |                                | 30                   |                                |
| Liczba godzin całkowitego nakładu pracy studenta (CNPS)   | 70                   |                                |                                | 90                   |                                |
| Forma zaliczenia  | Zaliczenie na ocenę* | Egzamin / zaliczenie na ocenę* | Egzamin / zaliczenie na ocenę* | Zaliczenie na ocenę* | Egzamin / zaliczenie na ocenę* |
| Dla grupy kursów zaznaczyć kurs końcowy (X)   |                      |                                |                                |                      |                                |
| Liczba punktów ECTS   | <b>2</b>             |                                |                                | <b>3</b>             |                                |
| w tym liczba punktów odpowiadająca zajęciom o charakterze praktycznym (P)                               | 0                    |                                |                                | 3                    |                                |
| w tym liczba punktów ECTS odpowiadająca zajęciom wymagającym bezpośredniego kontaktu (BK) (30+30 zajęć) | 1,2                  |                                |                                | 1,8                  |                                |

\*niepotrzebne skreślić

**WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I INNYCH KOMPETENCJI**

1. Wiedza z zakresu programowania
2. Podstawy sztucznej inteligencji
3. Umiejętność czytania ze zrozumieniem tekstów naukowych i technicznych w języku angielskim

**CELE PRZEDMIOTU**

C1 Uświadomienie studentom możliwości wykorzystania języka naturalnego jako narzędzia komunikacji i zapisu informacji w systemie informatycznym.

C2 Przedstawienie sposobów wykorzystania technologii językowych w systemach informacyjnych.

C3 Zapoznanie studentów z typami narzędzi i zasobów językowych i ich dostępnością ze szczególną uwagą poświęconą polskiej technologii językowej.

C4 Osiągnięcie przez studentów podstawowe umiejętności w zakresie formalnego opisu języka naturalnego oraz konstrukcji inteligentnych systemów przetwarzających wypowiedzi w języku naturalnym.

|  |
|--|
|  |
|--|

### **PRZEDMIOTOWE EFEKTY KSZTAŁCENIA**

Z zakresu wiedzy:

PEK\_W01 Posiada podstawową wiedzę w zakresie formalnego opisu języka naturalnego.

PEK\_W02 Zna metody konstrukcji inteligentnych systemów przetwarzających wypowiedzi w języku naturalnym.

PEK\_W03 Zna podstawowe zasoby i narzędzia językowe oraz zakres ich wykorzystania w ramach systemów informatycznych.

PEK\_W04 Zna stan bieżący technologii językowej dla języka angielskiego i polskiego.

Z zakresu umiejętności:

PEK\_U01 Umie zastosować metody inżynierii języka naturalnego w budowie systemów inteligentnego wyszukiwania informacji oraz praktycznych systemów wydobywania informacji z tekstu.

PEK\_U02 Zna biegle typy narzędzi językowych i ich dostępność, potrafi je odpowiednio zestawić i zastosować.

PEK\_U03 Umie zestawić podstawowe zasoby i narzędzia językowe w odpowiedni ciąg przetwarzający język naturalny.

PEK\_U04 Ma podstawowe umiejętności w zakresie formalnego opisu języka naturalnego oraz konstrukcji inteligentnych systemów przetwarzających wypowiedzi w języku naturalnym.

PEK\_U05 Umie odnaleźć w dostępnych zasobach internetowych informację na temat aktualnego stanu technologii językowych dla języka polskiego i angielskiego.

Z zakresu kompetencji społecznych:

PEK\_K01 Potrafi wyjaśniać znaczenie technologii językowej dla budowy systemów informatycznych.

PEK\_K02 Potrafi dobrać i skompletować interdyscyplinarny zespół do konstrukcji systemu przetwarzającego tekst w języku naturalnym.

### **TREŚCI PROGRAMOWE**

| <b>Forma zajęć - wykład</b> |   | <b>Liczba godzin</b> |
|-----------------------------|---|----------------------|
| Wy1                         | Podstawy lingwistyki informatycznej: języka naturalny jako narzędzie komunikacji, podstawowe poziomy opisu języka naturalnego, modele formalne w opisie języka naturalnego. | 2                    |
| Wy2                         | Charakterystyka typowych etapów przetwarzania języka naturalnego. pojęcia zasobów językowych i narzędzi językowych; przegląd zastosowań inżynierii języka naturalnego.      | 2                    |
| Wy3                         | Tokenizacja i segmentacja tekstu oraz podstawy technologii automatów skończone stanowych.   | 2                    |
| Wy4                         | Analiza morfosyntaktyczna: analizatory morfologiczne, transduktory, ujednolaczanie opisu, konstrukcja i zastosowanie tagerów.   | 2                    |
| Wy5                         | Analiza składniowa: wykorzystanie formalizmów gramatyki, głęboka  | 2                    |

|     |   |           |
|-----|---|-----------|
|     | oraz płytka, konstrukcja parserów.  |           |
| Wy6 | Reprezentacja znaczeń słów i wyrażeń: logiczna reprezentacja znaczenia, leksykalne sieci semantyczne (wordnets), rzutowanie na ontologie, połączone otwarte dane (Linked Open Data) | 2         |
| W7  | Analiza semantyczna: głęboka — oparta na reprezentacji semantycznej, oraz płytka — analiza poprzez cechy; ujednoznacznianie sensu słów; analiza dyskursu.                           | 2         |
| W8  | Wprowadzenie do wydobywania informacji. Środowiska programistyczne do przetwarzania języka naturalnego. Infrastruktury technologii językowych.                                      | 2         |
| W9  | Rozpoznawanie nazw własnych i wyrażeń identyfikujących. Rozpoznawanie powiązań koreferencyjnych.  | 2         |
| W10 | Rozpoznawanie i wydobywanie relacji semantycznych z tekstu. Rozpoznawanie odniesień do sytuacji w tekście. Agregowanie wydobytej informacji.  | 2         |
| W11 | Semantyka dystrybucyjna i jej zastosowania do nienadzorowanego uczenia się znaczeń słów, wyrażeń językowych i dokumentów.   | 2         |
| W12 | Klasyfikacja, grupowanie i filtrowanie dokumentów tekstowych.   | 2         |
| W13 | Systemy do odpowiadania na pytania w języku naturalnym.   | 2         |
| W14 | Systemy wielojęzyczne, międzyjęzykowe (cross-lingual); maszynowe tłumaczenie.   | 2         |
| W15 | Systemy dialogowe i systemy rozumienia tekstu.  | 2         |
|     | Suma godzin   | <b>30</b> |

| Forma zajęć - ćwiczenia |             | Liczba godzin |
|-------------------------|-------------|---------------|
| Ćw1                     |             |               |
| Ćw2                     |             |               |
| Ćw3                     |             |               |
| Ćw4                     |             |               |
| ..                      |             |               |
|                         | Suma godzin |               |

| Forma zajęć - laboratorium |             | Liczba godzin |
|----------------------------|-------------|---------------|
| La1                        |             |               |
| La2                        |             |               |
| La3                        |             |               |
| La4                        |             |               |
| La5                        |             |               |
| ...                        |             |               |
|                            | Suma godzin |               |

| Forma zajęć - projekt |   | Liczba godzin |
|-----------------------|---|---------------|
| Pr1                   | Wybór tematu oraz jego szczegółowe doprecyzowanie.                              | 2             |
| Pr2                   | Przegląd literatury oraz zapoznanie się z niezbędnymi podstawami teoretycznymi. | 4             |
| Pr3                   | Wybór metod do implementacji.   | 2             |
| Pr4                   | Wybór niezbędnych narzędzi i komponentów programistycznych oraz                 | 2             |

|     |   |           |
|-----|---|-----------|
|     | narzędzi i zasobów językowych.  |           |
| Pr5 | Instalacja wybranych komponentów programistycznych, zapoznanie się z ich wykorzystaniem, zintegrowanie ze stosowanym środowiskiem developerskim, dopasowanie elementów technologii językowej do konkretnego problemu. | 4         |
| Pr6 | Implementacja.  | 6         |
| Pr7 | Zgromadzenie niezbędnych danych do badań i eksperymentów.   | 4         |
| Pr8 | Przeprowadzenie eksperymentów i optymalizacja parametrów metod.   | 4         |
| Pr9 | Przygotowanie raportu końcowego.  | 2         |
|     | <b>Suma godzin</b>  | <b>30</b> |

| Forma zajęć - seminarium |             | Liczba godzin |
|--------------------------|-------------|---------------|
| Se1                      |             |               |
| Se2                      |             |               |
| Se3                      |             |               |
| ...                      |             |               |
|                          | Suma godzin |               |

| STOSOWANE NARZĘDZIA DYDAKTYCZNE  |
|--|
| <p>N1. Podręczniki.</p> <p>N2. Materiały elektroniczne na wskazanych stronach i serwisach internetowych.</p> <p>N3. Udostępnione zasoby i narzędzia językowe dla języka polskiego.</p> <p>N4 Zasoby i narzędzia językowe oraz podstawowe architektury przetwarzania języka naturalnego dostępne na wskazanych stronach internetowych.</p> <p>N5 Materiały do wykładu i projektu udostępnione poprzez portal E-learning Wydziału Informatyki i Zarządzania.</p> |

#### OCENA OSIĄGNIĘCIA PRZEDMIOTOWYCH EFEKTÓW KSZTAŁCENIA

| Oceny (F – formująca (w trakcie semestru), P – podsumowująca (na koniec semestru)) | Numer efektu kształcenia               | Sposób oceny osiągnięcia efektu kształcenia   |
|--|--|---|
| F1 Ocena fazy przygotowania projektu   | PEK_U02, PEK_U03, PEK_U05              | Student przedstawia opracowanie w formie raportu, który podlega ocenie.               |
| F2 Ocena konstrukcji systemu   | PEK_U01, PEK_U04                       | Student przedstawia opracowanie w formie raportu, który podlega ocenie.               |
| F3 Ocena kompleksowa systemu i wyników eksperymentów                               | PEK_U01-PEK_U05 oraz PEK_K01 i PEK_K02 | Student prezentuje osiągnięte wyniki i przedstawia końcowy raport podlegający ocenie. |
| P kolokwium zaliczeniowe – efekty PEK_W01- PEK_W04                                 |  |   |

## LITERATURA PODSTAWOWA I UZUPEŁNIAJĄCA

### **LITERATURA PODSTAWOWA:**

- [1] Handbook of Natural Language Processing (Second Edition). (Ed.) Nitin Indurkha i Fred J. Damerau. CRC Press, 2010
- [2] Sholom M. Weiss, Nitin Indurkha, Tong Zhang i Fred Damerau. Text Mining: Predictive Methods for Analyzing Unstructured Information, 2010.
- [3] Christopher D. Manning, Prabhakar Raghavan i Hinrich Schütze. Introduction to Information Retrieval. Cambridge Univ. Press, 2008.
- [4] Jurafsky, D. & Martin, J. H. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition Prentice Hall, 2000.
- [5] Manu Konchady Text Mining Application Programming (Programming Series) Charles River Media, Inc., 2006.
- [6] Mykowiecka A. Inżynieria lingwistyczna, Komputerowe przetwarzanie tekstów w języku naturalnym, Wydawnictwo PJWSTK, Warszawa 2007
- [7] Piasecki Maciej. Selektywne wprowadzenie do semantyki formalnej. Red. Szymanik J. i Zajenkowski M., Kognitywistyka. O umyśle umyślnie i nieumyślnie, Koło Filozoficzne przy MISH, Uniwersytet Warszawski, str. 113-155, 2004.
- [8] Marius Paşca. Open-Domain Question Answering from Large Text Collections. CSLI, Stanford, 2003.

### **LITERATURA UZUPEŁNIAJĄCA:**

- [1] Daniel Bikel i Imed Zitouni. Multilingual Natural Language Processing Applications: From Theory to Practice
- [2] Manning, C. D. i Schütze, H. Foundations of Statistical Natural Language Processing The MIT Press, 2001.
- [3] Mitkov, R. (ed.) The Oxford Handbook of Computational Linguistics Oxford University Press, 2003.
- [4] Piasecki, M.; Szpakowicz, S. & Broda, B. (2009), *A Wordnet from the Ground Up*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, URL : [http://www.plwordnet.pwr.wroc.pl/main/content/files/publications/A\\_Wordnet\\_from\\_the\\_Ground\\_Up.pdf](http://www.plwordnet.pwr.wroc.pl/main/content/files/publications/A_Wordnet_from_the_Ground_Up.pdf)

### **OPIEKUN PRZEDMIOTU (IMIE, NAZWISKO, ADRES E-MAIL)**

Maciej Piasecki, [maciej.piasecki@pwr.wroc.pl](mailto:maciej.piasecki@pwr.wroc.pl)

MACIERZ POWIĄZANIA EFEKTÓW KSZTAŁCENIA DLA PRZEDMIOTU  
*Inżynieria języka naturalnego*  
Z EFEKTAMI KSZTAŁCENIA NA KIERUNKU *Informatyka, II stopień, profil ogólnoakademicki*  
I SPECJALNOŚCI *Inteligentne Systemy Informatyczne*

| Przedmiotowy efekt kształcenia | Odniesienie przedmiotowego efektu do efektów kształcenia zdefiniowanych dla kierunku studiów i specjalności (o ile dotyczy)** | Cele przedmiotu*** | Treści programowe*** | Numer narzędzia dydaktycznego*** |
|--------------------------------|---|--------------------|----------------------|----------------------------------|
| <b>PEK_W01 (wiedza)</b>        | K2INF_W06_S2ISI_W04, K2INF_W02  | C4                 | Wy1                  | 1,2,5                            |
| <b>PEK_W02</b>                 | K2INF_W06_S2ISI_W04, K2INF_W02, K2INF_W05   | C2                 | Wy7                  | 1,2,5                            |
| <b>PEK_W03</b>                 | K2INF_W06_S2ISI_W04, K2INF_W02, K2INF_W05, K2INF_W07  | C3                 | Wy2-Wy6              | 1,2,3,4,5                        |
| <b>PEK_W04</b>                 | K2INF_W06_S2ISI_W04, K2INF_W02, K2INF_W05   | C1                 | Wy3-Wy6              | 1,2,3,4,5                        |
| <b>PEK_U01 (umiejętności)</b>  | K2INF_U08_S2ISI_U06, K2INF_U02, K2INF_U09   | C1                 | Pr1,Pr8,Pr9          | 1,2,3,4,5                        |
| <b>PEK_U02</b>                 | K2INF_U08_S2ISI_U08, K2INF_U09  | C3                 | Pr2                  | 1,2,3,4,5                        |
| <b>PEK_U03</b>                 | K2INF_U08_S2ISI_U06, K2INF_U05, K2INF_U06, K2INF_U07, K2INF_U09   | C2                 | Pr3-Pr6              | 1,2,3,4,5                        |
| <b>PEK_U04</b>                 | K2INF_U08_S2ISI_U09, K2INF_U09  | C4                 | Pr3,Pr4,Pr8          | 1,2,3,4,5                        |
| <b>PEK_U05</b>                 | K2INF_U08_S2ISI_U06, K2INF_U01, K2INF_U04, K2INF_U10, K2INF_U09   | C3                 | Pr7                  | 1,2,3,4,5                        |
| <b>PEK_K01 (kompetencje)</b>   | K2INF_U08_S2ISI_U06, K2INF_K01  | C1                 | Pr9                  | 1,2,5                            |
| <b>PEK_K02</b>                 | K2INF_U08_S2ISI_U06, K2INF_K01, K2INF_K03, K2INF_K05  | C4                 | Pr1,Pr9              | 1,2,5                            |

\*\* - wpisać symbole kierunkowych/specjalnościowych efektów kształcenia

\*\*\* - z tabeli powyżej