

SYLABUS – OPIS ZAJĘĆ/PRZEDMIOTU

I. Informacje ogólne

1. Nazwa zajęć/przedmiotu: **Przetwarzanie Języka Naturalnego**
2. Kod zajęć/przedmiotu: **PJN**
3. Rodzaj zajęć/przedmiotu (obowiązkowy lub fakultatywny): **obowiązkowy**
4. Kierunek studiów: **Język, Umysł, Technologia**
5. Poziom studiów (I lub II stopień, jednolite studia magisterskie): **II stopień**
6. Profil studiów (ogólnoakademicki / praktyczny): **ogólnoakademicki**
7. Rok studiów (jeśli obowiązuje): **drugi**
8. Rodzaje zajęć i liczba godzin: **30h W, 30h CW**
9. Liczba punktów ECTS: **6**
10. Imię, nazwisko, tytuł / stopień naukowy, adres e-mail prowadzącego zajęcia: **dr Rafał Jaworski, rjawor@amu.edu.pl**
11. Język wykładowy: **polski**
12. Zajęcia / przedmiot prowadzone zdalnie (e-learning) (tak [częściowo/w całości] / nie): **nie**

II. Informacje szczegółowe

1. Cele zajęć/przedmiotu:

Celem przedmiotu jest zapoznanie studentów z zagadnieniami automatycznego przetwarzania języka naturalnego przy użyciu oprogramowania komputerowego. W pierwszej części przedstawione zostaną podstawowe techniki analizy tekstu, takie jak podział na zdania, na wyrazy, normalizacja tekstu. Następnie omówione zostaną zagadnienia związane z głębszą analizą języka, np. parsing płytki i głęboki, algorytmy klasyfikacji tekstu wykorzystujące uczenie maszynowe oraz tłumaczenie automatyczne. Po ukończeniu przedmiotu studenci będą potrafili implementować elementarne techniki przetwarzania języka naturalnego oraz uruchamiać bardziej zaawansowane narzędzia.

2. Wymagania wstępne w zakresie wiedzy, umiejętności oraz kompetencji społecznych (jeśli obowiązują):

Przedmiot wymaga podstawowej wiedzy w zakresie algorytmiki i programowania.

3. Efekty uczenia się (EU) dla zajęć i odniesienie do efektów uczenia się (EK) dla kierunku studiów:

Symbol EU dla zajęć/przedmiotu	Po zakończeniu zajęć i potwierdzeniu osiągnięcia EU student/ka:	Symbole EK dla kierunku studiów
PJN_01	Zna podstawowe zastosowania przetwarzania języka naturalnego, zna granice możliwości technologicznych, potrafi wymienić podstawowe techniki przetwarzania języka.	K_W07 K_W09 K_W11 K_U02
PJN_02	Potrafi napisać wyrażenie regularne, stosuje wyrażenia regularne do praktycznych zadań z dziedziny przetwarzania tekstu, potrafi zastosować wyrażenia regularne w języku Python.	K_W11 K_U08 K_U14
PJN_03	Zna pojęcie gramatyki formalnej, potrafi zaprojektować własną gramatykę oraz zmodyfikować istniejącą, implementuje gramatykę bezkontekstową w języku Python.	K_W11 K_U08 K_U14
PJN_04	Potrafi uruchomić powłokę systemu Linux, zna podstawowe komendy do przetwarzania tekstu pod Linuxem, rozumie koncepcję potoku komend.	K_W11 K_U08 K_U14
PJN_05	Zna techniki dzielenia tekstu na zdania oraz wyrazy, rozumie wiążące się z tym wyzwania (zwłaszcza w językach, które nie stosują spacji), potrafi zdefiniować reguły podziału na zdania i wyrazy.	K_W11 K_U08
PJN_06	Wie, czym jest analiza morfologiczna tekstu, zna elektroniczne słowniki morfologiczne dla różnych języków, zna techniki dezambiguacji.	K_W11 K_W13 K_U08
PJN_07	Zna najpopularniejsze toolkity do przetwarzania języka naturalnego, potrafi je zainstalować i uruchomić.	K_W11 K_U08
PJN_08	Zna zastosowanie technik uczenia maszynowego w przetwarzaniu języka, potrafi zaimplementować klasyfikator Bayesa używając gotowej implementacji algorytmu.	K_W11 K_U08 K_U16

PJN_09	Zna techniki regresji statystycznej i ich zastosowanie w przetwarzaniu języka naturalnego.	K_W11 K_U08 K_U16
PJN_10	Rozumie pojęcie i zna zastosowanie sieci neuronowych w modelowaniu języka oraz problemach klasyfikacyjnych.	K_W11 K_W13 K_U08 K_U14
PJN_11	Zna statystyczne metody modelowania języka, potrafi opracować model dowolnego języka w oparciu o zasoby tekstowe.	K_W11 K_U08 K_U16
PJN_12	Zna techniki automatycznego sprawdzania pisowni, potrafi uruchomić korektor pisowni.	K_W11 K_U07 K_U08
PJN_13	Rozumie pojęcie parsingu płytkiego, potrafi uruchomić parser płytki.	K_W11 K_W13 K_U08
PJN_14	Zna pojęcie parsingu głębokiego i analizy składniowej. Potrafi uruchomić parser głęboki i analizować wyniki jego działania.	K_W11 K_W13 K_U08
PJN_15	Zna techniki automatycznego pozyskiwania tekstu i danych z Internetu, potrafi zaimplementować prosty web crawler.	K_W11 K_U08
PJN_16	Jest gotowy do współpracy z potencjalnym pracodawcą.	K_W08 K_K05 K_K11

4. Treści programowe zapewniające uzyskanie efektów uczenia się (EU) z odniesieniem do odpowiednich efektów uczenia się (EU) dla zajęć/przedmiotu

Treści programowe dla zajęć/przedmiotu:	Symbol EU dla zajęć/przedmiotu
Zastosowania przetwarzania języka naturalnego, podstawowa terminologia związana z tą dziedziną, granice możliwości PJN, uruchamianie programów PJN.	PJN_01
Wyrażenia regularne – zdefiniowanie języka wyrażeń regularnych, wykorzystanie do wyszukiwania i modyfikacji tekstu.	PJN_02
Gramatyki i języki formalne, budowanie własnej gramatyki kontekstowej, wykorzystanie gramatyk formalnych do analizy tekstu.	PJN_03
Polecenia systemu Linux, pojęcie potoku, zastosowanie w przetwarzaniu tekstu.	PJN_04
Segmentacja tekstu – dzielenie na zdania (segmenty) oraz słowa, formalizm SRX, algorytmy tokenizacji.	PJN_05
Analiza morfologiczna – słowniki morfologiczne, problem dezambiguacji, uruchamianie narzędzi do analizy morfologicznej.	PJN_06
Toolkity PJN – omówienie i zastosowanie najpopularniejszych toolkitów PJN dostępnych dla języka Python: NLTK oraz spaCy.	PJN_07
Wprowadzenie do uczenia maszynowego, zastosowanie uczenia maszynowego w przetwarzaniu języka, klasyfikator Bayesa.	PJN_08
Metody regresji linowej i logistycznej w przetwarzaniu języka naturalnego wraz z przykładami praktycznego ich zastosowania.	PJN_09
Koncepcja sieci neuronowych, ich znaczenie i zastosowanie w przetwarzaniu języka naturalnego.	PJN_10
Statystyczne metody modelowania języka i ich zastosowanie praktyczne.	PJN_11
Problem automatycznej korekty pisowni, zastosowanie programów do korekty pisowni w praktyce, wydajność algorytmów	PJN_12
Parsing płytki jako wydajna metoda powierzchniowej analizy tekstu wraz z zastosowaniami.	PJN_13
Parsing głęboki, różnice w stosunku do parsingu płytkiego, zastosowanie oraz przykłady.	PJN_14
Zautomatyzowane pobieranie zasobów lingwistycznych z Internetu, ćwiczenia praktyczne.	PJN_15
Uczestnictwo w wizycie studyjnej.	PJN_16

5. Zalecana literatura:

- Lane Hobson, Cole Howard, Hannes Hapke: „Przetwarzanie języka naturalnego w akcji”
- Edward Loper, Steven Bird, Ewan Klein: “Natural Language Processing with Python”

III. Informacje dodatkowe

1. Metody i formy prowadzenia zajęć umożliwiające osiągnięcie założonych EU (proszę wskazać z proponowanych metod właściwe dla opisywanych zajęć lub/i zaproponować inne)

Metody i formy prowadzenia zajęć	X
Wykład z prezentacją multimedialną wybranych zagadnień	X
Wykład konwersatoryjny	
Wykład problemowy	
Dyskusja	X
Praca z tekstem	X
Metoda analizy przypadków	
Uczenie problemowe (Problem-based learning)	X
Gra dydaktyczna/symulacyjna	
Rozwiązywanie zadań (np.: obliczeniowych, artystycznych, praktycznych)	X
Metoda ćwiczeniowa	X
Metoda laboratoryjna	X
Metoda badawcza (dociekania naukowego)	
Metoda warsztatowa	
Metoda projektu	X
Pokaz i obserwacja	
Demonstracje dźwiękowe i/lub video	
Metody aktywizujące (np.: „burza mózgów”, technika analizy SWOT, technika drzewka decyzyjnego, metoda „kuli śniegowej”, konstruowanie „map myśli”)	
Praca w grupach	
Inne (jakie?) -	
...	

2. Sposoby oceniania stopnia osiągnięcia EU (proszę wskazać z proponowanych sposobów właściwe dla danego EU lub/i zaproponować inne)

Sposoby oceniania	Symbole EU dla zajęć/przedmiotu				
	PJN_01-15	PJN_01-06	PJN_07-15		
Egzamin pisemny					
Egzamin ustny					
Egzamin z „otwartą książką”					
Kolokwium pisemne	x				
Kolokwium ustne					
Test					
Projekt			x		
Esej					
Raport		x			
Prezentacja multimedialna					
Egzamin praktyczny (obserwacja wykonawstwa)					
Portfolio					
Inne (jakie?) -					
...					

3. Nakład pracy studenta i punkty ECTS

Forma aktywności		Średnia liczba godzin na zrealizowanie aktywności
Godziny zajęć (wg planu studiów) z nauczycielem		60
Praca własna studenta *	Przygotowanie do zajęć	20
	Czytanie wskazanej literatury	17
	Przygotowanie pracy pisemnej, raportu, prezentacji, demonstracji, itp.	25
	Przygotowanie projektu	
	Przygotowanie pracy semestralnej	
	Przygotowanie do egzaminu / zaliczenia	20
	Inne (jakie?) – Wizyta studyjna	8
...		
SUMA GODZIN		150
LICZBA PUNKTÓW ECTS DLA ZAJĘĆ/PRZEDMIOTU		6

* proszę wskazać z proponowanych przykładów pracy własnej studenta właściwe dla opisywanych zajęć lub/i zaproponować inne

4. Kryteria oceniania wg skali stosowanej w UAM:

bardzo dobry (bdb; 5,0):

dobry plus (+db; 4,5):

dobry (db; 4,0):

dostateczny plus (+dst; 3,5):

dostateczny (dst; 3,0):

niedostateczny (ndst; 2,0):