

## OPIS ZAJĘĆ/PRZEDMIOTU (SYLABUS)

### I. Informacje ogólne

1. Nazwa zajęć/przedmiotu – **Elective 3: Natural Language Processing / Linguistic epistemology / Lexicography**
2. Kod zajęć/przedmiotu – 15-EL3-EL-11
3. Rodzaj zajęć/przedmiotu (obowiązkowy lub fakultatywny) – obowiązkowy
4. Kierunek studiów – English Linguistics: Theories, Interfaces, Technologies
5. Poziom studiów (I lub II stopień, jednolite studia magisterskie) – poziom I
6. Profil studiów (ogólnoakademicki / praktyczny) – ogólnoakademicki
7. Rok studiów (jeśli obowiązuje) – 3 rok
8. Rodzaje zajęć i liczba godzin (np.: 15 h W, 30 h ĆW) – 30 h ćw.
9. Liczba punktów ECTS – 4
10. Imię, nazwisko, tytuł / stopień naukowy, adres e-mail wykładowcy (wykładowców\*) / prowadzących zajęcia – dr Grzegorz Krynicki
11. Język wykładowy – angielski
12. Moduł zajęć / przedmiotu prowadzony zdalnie (e-learning) (tak [częściowo/w całości] / nie) – tak, w całości

\*proszę podkreślić koordynatora przedmiotu

### II. Informacje szczegółowe

#### 1. Cele zajęć/przedmiotu

- Kontynuacja kursu *IT Skills for Linguists 2*
- Wykorzystania średnio-zaawansowanych technik programowania w języku Python do zadań związanych z przetwarzaniem danych językowych przy pomocy biblioteki NLTK na poziomie podstawowym
- Stworzenie przez Studenta w ramach projektu zaliczeniowego oprogramowania związanego z tematem pracy licencjackiej.

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

- Umiejętność wykorzystania komputera, znajomość wyrażeń regularnych (*IT Skills for Linguists 1*)
- Znajomość języka Python na poziomie podstawowym (*IT Skills for Linguists 2*)
- Znajomość zagadnień z zakresu językoznawstwa (*Linguistic theory: Semantics, Linguistic theory: Syntax*)
- Znajomość podstaw składni i morfologii języka naturalnego (*Linguistic theory: Syntax*)
- Znajomość języka angielskiego na poziomie B1

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

Symbol EU dla modułu zajęć/przedmiotu	Po zakończeniu modułu i potwierdzeniu osiągnięcia EK student /ka:	Symbole EK dla kierunku studiów
EU_01	Potrąfi zidentyfikować obszary wykorzystania języków programowania w dziedzinie językoznawstwa	K_W02, K_W01, K_W06, K_U03, K_U09, K_K06
EU_02	Potrąfi dobrać język programowania i biblioteki stosownie do problemu	K_U02, K_U03, K_U05, K_U07, K_U09
EU_03	Potrąfi zaprojektować algorytm automatyzujący rozwiązanie problemu	K_U04, K_U05, K_U09, K_U15



EU_04	Potrafi sformułować struktury danych oraz składnię programu dla realizacji algorytmu	K_U09
EU_05	Potrafi odnajdywać źródła poszerzające jego wiedzę konieczną do rozwiązania problemu	K_U02, K_K01, K_K03
EU_06	Potrafi współpracować z członkami grupy w celu realizacji zadania programistycznego	K_U01, K_U10, K_U11, K_U12
EU_07	Potrafi szacować skuteczność stworzonego oprogramowania, jego optymalizacji i dokumentacji	K_U08

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/symbole EU dla zajęć / przedmiotu
Korzystanie z interpreterów/kompilatorów języka Python	EU_02, EU_04
Składnia i biblioteki języka Python	EU_03, EU_04
Obszary zastosowania języków programowania w przetwarzaniu języka naturalnego	EU_01, EU_02, EU_05, EU_06
Struktury danych i algorytmy wykorzystywane do przetwarzaniu języka naturalnego	EU_03, EU_04, EU_05, EU_06
Wstępne przetwarzanie tekstu, formaty przechowywania danych językowych	EU_03, EU_04, EU_07, EU_05, EU_06
Identyfikacja nazw własnych, kolokacji i n-gramów; oznaczanie części mowy	EU_03, EU_04, EU_07, EU_05, EU_06
Automaty skończenie stanowe, gramatyki formalne, wyrażenia regularne, budowa gramatyki opartej na atrybutach (Feature Based Grammar)	EU_03, EU_04, EU_07, EU_05, EU_06
Dostępne biblioteki programistyczne dla przetwarzania danych językowych	EU_01, EU_02, EU_03
Miary jakości wyników zwracanych przez stworzone oprogramowanie	EU_01, EU_07

5. Zalecana literatura:

- Bird, Steven - Ewan Klein - Edward Loper. Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit. URL: <http://www.nltk.org/book/>
- Christiansen, H. 2010. Logic Programming for Linguistics. Triangle: Tarragona. URL: <https://revistes.urv.cat/index.php/triangle/article/view/2376/2291>
- Clocksin, W.F. - C.S. Mellish.. 2003. Programming in Prolog (5<sup>th</sup> ed.). Springer-Verlag: Berlin Heidelberg.
- Hammond, Michael. 2017. Python for Linguists, manuscript.
- Nugues, Pierre M. 2014. Language Processing with Perl and Prolog - Theories, Implementation, and Application 2nd ed. Springer-Verlag: Berlin Heidelberg.
- Perkins, Jacob. 2014. Python 3 Text Processing with NLTK 3 Cookbook (2<sup>nd</sup> ed.). Birmingham: Packt Publishing.

### 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	
1. Wykład z prezentacją multimedialną wybranych zagadnień	✓
2. Wykład konwersatoryjny	✓
3. Wykład problemowy	
4. Dyskusja	✓
5. Praca z tekstem	✓
6. Metoda analizy przypadków	
7. Uczenie problemowe (Problem-based learning)	✓
8. Gra dydaktyczna/symulacyjna	
9. Rozwiązywanie zadań (np.: obliczeniowych, artystycznych, praktycznych)	✓
10. Metoda ćwiczeniowa	✓
11. Metoda laboratoryjna	✓
12. Metoda badawcza (dociekania naukowego)	
13. Metoda warsztatowa	✓
14. Metoda projektu	✓
15. Pokaz i obserwacja (prezentacje studenckie)	✓
16. Demonstracje dźwiękowe i/lub video	✓
17. Metody aktywizujące (np.: „burza mózgów”, technika analizy SWOT, technika drzewka decyzyjnego, metoda „kuli śniegowej”, konstruowanie „map myśli”)	✓
18. Praca w grupach	✓
19. Indywidualne konsultacje z prowadzącym	✓
20. Zadania na platformie e-learningowej Moodle, <a href="https://www.onlinegdb.com">https://www.onlinegdb.com</a> , <a href="https://codeboard.io">https://codeboard.io</a> , <a href="https://repl.it">https://repl.it</a>	✓

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

Sposoby oceniania	Symbole EK dla modułu zajęć/przedmiotu						
	EU_01	EU_02	EU_03	EU_04	EU_05	EU_06	EU_07
Egzamin pisemny							
Egzamin ustny							
Egzamin z „otwartą książką”							
Kolokwium pisemne							
Kolokwium ustne							
Test			✓				✓
Projekt	✓	✓	✓	✓	✓	✓	✓
Esej	✓	✓					
Raport	✓	✓	✓	✓	✓	✓	✓
Prezentacja multimedialna							
Egzamin praktyczny (obserwacja wykonawstwa)							
Portfolio							
Zadanie domowe wykonania program komputerowego		✓	✓	✓	✓		✓



Zadanie klasowe wykonania programu komputerowego		✓	✓	✓			✓
--	--	---	---	---	--	--	---

### 3. Nakład pracy studenta i punkty ECTS

Forma aktywności		Średnia liczba godzin na zrealizowanie aktywności
Godziny zajęć (wg planu studiów) z nauczycielem		30
Praca własna studenta*	Przygotowanie do zajęć	5
	Czytanie wskazanej literatury	5
	Przygotowanie raportu i prezentacji	10
	Przygotowanie projektu	10
	Przygotowanie pracy semestralnej	
	Przygotowanie do egzaminu / zaliczenia	
	...	
SUMA GODZIN		60
LICZBA PUNKTÓW ECTS DLA MODUŁU ZAJĘĆ/PRZEDMIOTU		4

\* proszę wskazać z proponowanych przykładów pracy własnej studenta właściwe dla opisywanego modułu lub/i zaproponować inne

### 4. Kryteria oceniania wg skali stosowanej w UAM:

- Ocena będzie średnią ważoną następujących składowych
  - Przygotowanie do zaj. (10%),
  - Zadania klasowe (10%),
  - Zadania domowe (20%),
  - Duże testy (30%),
  - Projekt (30%)
- Średnia posłuży do wyznaczenia oceny wg następującej skali
  - bardzo dobry (bdb; 5,0): znakomita wiedza i umiejętności
  - dobry plus (+db; 4,5): bardzo dobra wiedza i umiejętności
  - dobry (db; 4,0): dobra wiedza i umiejętności
  - dostateczny plus (+dst; 3,5): dobra wiedza i umiejętności, ale z niedociągnięciami
  - dostateczny (dst; 3,0): zadowalająca wiedza i umiejętności
  - niedostateczny (ndst; 2,0): niezadowalająca wiedza i umiejętności



## SYLLABUS – A COURSE DESCRIPTION

### I. General information

1. Course name: **Elective 3: Natural Language Processing / Linguistic epistemology / Lexicography**
2. Course code: 15-EL3-EL-11
3. Course type (compulsory or optional): compulsory
4. Study programme name: English Linguistics: Theories, Interfaces, Technologies
5. Cycle of studies: 1st cycle
6. Educational profile (general academic profile or practical profile): academic
7. Year of studies (if relevant): 3
8. Type of classes and number of contact hours: tutorial 30 h
9. Number of ECTS credits: 4
10. Name, surname, academic degree/title, and e-mail address of the module lecturer (lecturers\*) / teachers: dr Grzegorz Krynicki
11. Language of instruction: English
12. Online learning (yes (partially / entirely) / no) – yes, entirely

\*please underline module coordinator's name

### II. Detailed information

1. Course aim (aims)
  - Introduce students to Natural Language Processing
  - Use Python programs to help you manipulate and analyse language data
  - Presenting how key concepts from NLP and linguistics are used to describe and analyse language
  - Learning about data structures and algorithms and their use in NLP
  - Learning how language data is stored in standard formats, and how data can be used to evaluate the performance of NLP techniques
2. Pre-requisites in terms of knowledge, skills and social competences (if relevant)
  - Using file managers, text editors, regular expression (*IT Skills for Linguists 1*)
  - Basic knowledge of Python (*IT Skills for Linguists 2*)
  - Basic knowledge of linguistics (*Linguistic theory: Semantics, Linguistic theory: Syntax*)
  - Basic knowledge of morphology and syntax (*Linguistic theory: Syntax*)
  - English at B1
3. Course learning outcomes (EU) in terms of knowledge, skills and social competences and their reference to study programme learning outcome

Course learning outcome symbol (EU)	On successful completion of the course and validation of its learning outcomes, a student:	Reference to study programme learning outcomes
EU_01	Students are able to identify areas of use of programming languages in the field of linguistics	K_W02, K_W01, K_W06, K_U03, K_U09, K_K06



EU_02	They can select the programming language, libraries and methods to solve the problem	K_U02, K_U03, K_U05, K_U07, K_U09
EU_03	They can design an algorithm to automate a solution to a problem	K_U04, K_U05, K_U09, K_U15
EU_04	They are able to formulate data structures and the program syntax for the implementation of the algorithm	K_U09
EU_05	They are able to find sources that broaden the knowledge necessary to solve a problem	K_U02, K_K01, K_K03
EU_06	They can work with the group to accomplish a programming task	K_U01, K_U10, K_U11, K_U12
EU_07	They are able to estimate the effectiveness of the created software, optimize it and document it	K_U08

#### 4. Learning content with reference to course learning outcomes (EU)

Course learning content	Course learning outcome symbol(s) (EU)
Using Python interpreters/compiler	EU_02, EU_04
Python syntax and libraries	EU_03, EU_04
Areas of application of programming languages in natural language processing	EU_01, EU_02, EU_05, EU_06
Data structures and algorithms used for natural language processing	EU_03, EU_04, EU_05, EU_06
Text pre-processing, language data storage formats	EU_03, EU_04, EU_07, EU_05, EU_06
Identification of proper names, collocation and n-grams; marking of parts of speech	EU_03, EU_04, EU_07, EU_05, EU_06
State completion automats, formal grammars, regular expressions, Feature Based Grammar	EU_03, EU_04, EU_07, EU_05, EU_06
Programming libraries available for language data processing	EU_01, EU_02, EU_03
Measures of quality of results returned by the developed software	EU_01, EU_07

#### 5. Reading list



- Bird, Steven - Ewan Klein - Edward Loper. Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit. URL: <http://www.nltk.org/book/>
- Christiansen, H. 2010. Logic Programming for Linguistics. Triangle: Tarragona. URL: <https://revistes.urv.cat/index.php/triangle/article/view/2376/2291>
- Clocksin, W.F. - C.S. Mellish.. 2003. Programming in Prolog (5<sup>th</sup> ed.). Springer-Verlag: Berlin Heidelberg.
- Hammond, Michael. 2017. Python for Linguists, manuscript.
- Nugues, Pierre M. 2014. Language Processing with Perl and Prolog - Theories, Implementation, and Application 2nd ed. Springer-Verlag: Berlin Heidelberg.
- Perkins, Jacob. 2014. Python 3 Text Processing with NLTK 3 Cookbook (2<sup>nd</sup> ed.). Birmingham: Packt Publishing.

### III. Additional information

1. Teaching and learning methods and activities to enable students to achieve the intended course learning outcomes (please indicate the appropriate methods and activities with a tick and/or suggest other methods.)

Teaching methods and techniques	
1. Lecture with a multimedia presentation	✓
2. Interactive lecture	✓
3. Problem-based lecture	
4. Class discussion	✓
5. Text-based work	✓
6. Case study work	
7. Problem-based learning	✓
8. Educational simulation / game	
9. Problem solving (e.g.: computational, artistic, practical)	✓
10. Practical exercises	✓
11. Laboratory practice	✓
12. Scientific inquiry method	
13. Workshop method	✓
14. Project work	✓
15. Presentation and observation (student presentations)	✓
16. Audio / Video demonstration	✓
17. Creative methods (e.g.: „brainstorming”, SWOT, decision tree, snowballing, mind mapping”)	✓
18. Group work	✓
19. Individuals consultations with teacher	✓
20. Moodle platform resources, <a href="https://www.onlinegdb.com">https://www.onlinegdb.com</a> , <a href="https://codeboard.io">https://codeboard.io</a> , <a href="https://repl.it">https://repl.it</a>	✓



2. Assessment methods to test if learning outcomes have been achieved (please indicate with a tick the appropriate methods for each LO (EU) and/or suggest different methods).

Assessment methods	Course learning outcome symbol						
	EU_01	EU_02	EU_03	EU_04	EU_05	EU_06	EU_07
Written exam							
Oral exam							
„Open book” exam							
Written final test (kolokwium)							
Oral test (kolokwium)							
Test (quiz)			✓				✓
Project	✓	✓	✓	✓	✓	✓	✓
Essay	✓	✓					
Report	✓	✓	✓	✓	✓	✓	✓
Multimedia presentation							
Practical exam (performance observation)							
Portfolio							
Other – Corrections of BA paper excerpts		✓	✓	✓	✓		✓
...		✓	✓	✓			✓

3. Student workload (ECTS credits)

Activity types		Mean number of hours spent on each activity type
Contact hours with the teacher as specified in the study programme		30
Students' self-study*	Preparation for classes	5
	Reading for classes	5
	Preparation of the report	10
	Preparation of the project	10
TOTAL HOURS		60
Total ECTS credits for the course		4

\* please indicate the appropriate activity types and/or propose different activities



4. Assessment criteria in accordance with AMU grading scale:

- quiz results (10%),
- class assignments (30%),
- home assignments (30%), as well as
- group project report (15%) and
- group project class presentation (15%)

The grade will be calculated from the mean according to the following scale:

- $\geq 92\%$  very good (bdb; 5,0): excellent knowledge and skills
- $\geq 84\%$  good plus (+db; 4,5): very good knowledge and skills
- $\geq 76\%$  good (db; 4,0): good knowledge and skills
- $\geq 68\%$  sufficient plus (+dst; 3,5): knowledge and skills with minor errors
- $\geq 60\%$  sufficient (dst; 3,0): knowledge and skills with major errors
- $< 60\%$  insufficient (ndst; 2,0): unsatisfactory knowledge and skills