



## Subject card

Subject name and code	Natural Language Processing, PG_00048268						
Field of study	Informatics						
Date of commencement of studies	February 2023		Academic year of realisation of subject		2023/2024		
Education level	second-cycle studies		Subject group		Optional subject group Subject group related to scientific research in the field of study		
Mode of study	Full-time studies		Mode of delivery		at the university		
Year of study	2		Language of instruction		Polish		
Semester of study	3		ECTS credits		2.0		
Learning profile	general academic profile		Assessment form		assessment		
Conducting unit	Department of Intelligent Interactive Systems -> Faculty of Electronics, Telecommunications and Informatics						
Name and surname of lecturer (lecturers)	Subject supervisor		dr hab. inż. Jan Daciuk				
	Teachers		dr hab. inż. Jan Daciuk				
Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	0.0	15.0	0.0	30
	E-learning hours included: 0.0						
Learning activity and number of study hours	Learning activity	Participation in didactic classes included in study plan		Participation in consultation hours		Self-study	SUM
	Number of study hours	30		4.0		16.0	50
Subject objectives	The purpose of this subject is to familiarize students with basic techniques used in natural language processing and to to prepare them for work on development of applications for that domain.						

Learning outcomes	Course outcome	Subject outcome	Method of verification
	[K7_U04] can apply knowledge of programming methods and techniques as well as select and apply appropriate programming methods and tools in computer software development or programming devices or controllers using microprocessors or programmable elements or systems specific to the field of study, making assessment and critical analysis of the prepared software as well as a synthesis and creative interpretation of information presented with it	Can use acquired knowledge to develop natural language processing software.	[SU4] Assessment of ability to use methods and tools [SU1] Assessment of task fulfilment
	[K7_W41] Knows and understands, to an increased extent, the standards, production methods, life cycle and development trends of software as well as information systems and applications.	Knows and profoundly understands standards, development methods, life cycle, and development trends of natural language processing software.	[SW1] Assessment of factual knowledge
	[K7_W03] Knows and understands, to an increased extent, the construction and operating principles of components and systems related to the field of study, including theories, methods and complex relationships between them and selected specific issues - appropriate for the curriculum.	Knows and understands principles and methods for building dictionaries, morphological analysis and synthesis, part-of-speech tagging, parsing and semantic analysis.	[SW1] Assessment of factual knowledge
	[K7_U06] can analyse the operation of components, circuits and systems related to the field of study; measure their parameters; examine technical specifications; interpret obtained results and draw conclusions	Can analyze functioning of natural language processing systems and can draw conclusions from that analysis.	[SU4] Assessment of ability to use methods and tools [SU2] Assessment of ability to analyse information
	[K7_U42] can solve engineering and research problems including design, assessment and maintenance of information systems and applications, using experimental methods and management techniques	Can solve engineering and research problems in design, evaluation, and maintenance of systems and application in natural language processing domain using experimental methods and management techniques.	[SU4] Assessment of ability to use methods and tools [SU3] Assessment of ability to use knowledge gained from the subject [SU2] Assessment of ability to analyse information [SU1] Assessment of task fulfilment
Subject contents	<b>Lecture:</b> <ol style="list-style-type: none"> <li>1. Introduction, segmentation.</li> <li>2. Morphology: concatenation.</li> <li>3. Morphology: spelling rules.</li> <li>4. Morphological analysis, lexicon as an automaton.</li> <li>5. Incremental construction of lexicons, perfect hashing.</li> <li>6. Spelling correction with automata.</li> <li>7. Spelling correction: noisy channel model.</li> <li>8. Document retrieval.</li> <li>9. Tagging.</li> <li>10. Syntax description methods.</li> <li>11. Parsing with unification.</li> <li>12. Representation of meaning.</li> <li>13. Syntax-driven semantic analysis.</li> <li>14. Dialogue.</li> <li>15. Word sense disambiguation.</li> </ol> <b>Project:</b> <ol style="list-style-type: none"> <li>1. Linux/Unix tools for text processing.</li> <li>2. Morphology.</li> <li>3. Segmentation.</li> <li>4. Document retrieval.</li> <li>5. Syntax and semantics.</li> </ol>		
Prerequisites and co-requisites	Programming skills (mainly arbitrary scripting languages), understanding Prolog programs.		

Assessment methods and criteria	Subject passing criteria	Passing threshold	Percentage of the final grade
	exam	50.0%	50.0%
	project	50.0%	50.0%
Recommended reading	Basic literature	<div>1. Daniel Jurafsky, James Martin, Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Second Edition, Prentice Hall, 2008.</div> <div>2. Christopher D. Manning, Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press, 2000.</div> <div>3. Emmanuel Roche, Yves Schabes, Finite-State Language Processing, MIT Press, 1997.</div> <div>4. Computational Linguistics journal and proceedings of ACL (Association for Computational Linguistics) conferences. Available from <a href="http://acl.ldc.upenn.edu/">http://acl.ldc.upenn.edu/</a> – ACL Anthology.</div>	
	Supplementary literature	<div>1. Zygmunt Saloni, Włodzimierz Gruszczyński, Marcin Woliński, Robert Wołosz, Słownik gramatyczny języka polskiego. Podstawy teoretyczne. Instrukcja użytkowania, Wiedza Powszechna, 2007.</div> <div>2. Stanisław Mędak, Słownik form koniugacyjnych czasowników polskich, Universitas, Kraków, 2004.</div> <div>3. Stanisław Mędak, Słownik odmiany rzeczowników polskich, Universitas, Kraków, 2003.</div>	
	eResources addresses	Adresy na platformie eNauczanie: Przetwarzanie języka naturalnego -2024 - Moodle ID: 36654 <a href="https://enauczanie.pg.edu.pl/moodle/course/view.php?id=36654">https://enauczanie.pg.edu.pl/moodle/course/view.php?id=36654</a>	
Example issues/ example questions/ tasks being completed			
Work placement	Not applicable		