SYLLABUS

1. Information regarding the programme

1.1 Higher education institution	**Babeş-Bolyai University**
1.2 Faculty	**Faculty of Mathematics and Computer Science**
1.3 Department	**Department of Computer Science**
1.4 Field of study	**Computer Science**
1.5 Study cycle	**Bachelor**
1.6 Study programme / Qualification	**Computer Science - english**

2. Information regarding the discipline

2.1 Name of the discipline			**Introduction to Natural Language Processing**				
2.2 Course coordinator			**Lecturer Ph.D. Lupea Mihaiela-Ana**				
2.3 Seminar coordinator			**Lecturer Ph.D. Lupea Mihaiela-Ana**				
2.4. Year of study	**3**	2.5 Semester	**6**	2.6. Type of evaluation	**C**	2.7 Type of discipline	**optional**

3. Total estimated time (hours/semester of didactic activities)

3.1 Hours per week	4	Of which: 3.2 course	2	3.3 seminar/laboratory	1 lab+ 1 pr
3.4 Total hours in the curriculum	48	Of which: 3.5 course	24	3.6 seminar/laboratory	24
Time allotment:					hours
Learning using manual, course support, bibliography, course notes					15
Additional documentation (in libraries, on electronic platforms, field documentation)					15
Preparation for seminars/labs, homework, papers, portfolios and essays					15
Tutorship					7
Evaluations					10
Practical project					15
3.7 Total individual study hours	77				
3.8 Total hours per semester	125				
3.9 Number of ECTS credits	5				

4. Prerequisites (if necessary)

4.1. curriculum	Formal languages, Data structures, Graphs Algorithms
4.2. competencies	Programming skills in a high level programming language

5. Conditions (if necessary)

5.1. for the course	
5.2. for the seminar /lab activities	• Laboratory with computers; high level programming language environment (.NET or any Java environment a.s.o.)

6. Specific competencies acquired

Professional competencies	• Assimilation of mathematical concepts and formal models to understand, verify and validate software systems ; • Advanced ability to approach, model and solve phenomena and problems from natural language and economy using fundamental knowledge from mathematics and computer science; • Ability to approach and solve complex problems using various techniques of computational intelligence; • Proficient use of methodologies and tools specific to programming languages and software systems.
Transversal competencies	• Etic and fair behavior, committment to professional deontology • Team work capabilities; able to fulfill different roles • Professional communication skills; concise and precise description, both oral and written, of professional results, negociation abilities; • Antepreneurial skills; working with economical knowledge; continuous learning • Good English communication skills

7. Objectives of the discipline (outcome of the acquired competencies)

7.1 General objective of the discipline	• To introduce the basic principles, domains and tasks in Natural Language Processing (NLP) • To understand the current state of the art in order to realize an overview of a specific domain in NLP and to implement a NLP tool.
7.2 Specific objective of the discipline	• Apply and use formal models (logics, grammars, parsing), statistic models (HMM), artificial intelligence algorithms and techniques to solve different tasks at the syntactic level (POS-tagging, parsing, chunking), and semantic level (keyword extraction, document summarization, anaphora resolution, sentiment analysis, word sense disambiguation) in Natural Language Processing domain.

8. Content

8.1 Course	Teaching methods	Remarks
Course 1. Natural Language Processing (NLP): stages, domains, applications.	Exposure: description, explanation, examples, debate, dialogue	
Course 2. **Part of speech tagging** WordNet and RoWordNet - knowledge structure, semantic relations, lexical relations	Exposure: description, explanation, examples, debate, dialogue	
Course 3. Text representation	Exposure: description, explanation, examples, dialogue	
Course 4. Syntactic parsing - grammar rules for English - sentence level construction; - Cocke-Kasami-Yonger (CKY) algorithm;	Exposure: description, explanation, dialogue, examples.	

Course 5. **Students' presentations of a NLP task/tool**	Debate, dialogue	
Course 6. Hidden Markov Model - Markov chains, Hidden Markov Model(HMM); - three canonical problems associated with HMM - the forward algorithm; Viterbi algorithm.	Exposure: description, explanation, examples, debate, dialogue	
Course 7. Keyword extraction - TextRank and RAKE algorithms	Exposure: description, explanation, examples, dialogue	
Course 8. Document summarization - approaches based on clustering, graphs.	Exposure: description, explanation, examples, dialogue	
Course 9. Sentiment analysis - opinion mining in social media - emotion analysis in literature	Exposure: description, explanation, debate, examples, dialogue	
Course 10. Anaphora resolution - Lapin and Lease algorithm - Mitkov's algorithm	Exposure: description, explanation, debate, examples, dialogue	
Course 11. Word Sense Disambiguation - dictionary and graph-based approaches.	Exposure: description, explanation, dialog, examples	
Course 12. **Students' presentations of the practical projects**	Debate, dialogue	

Bibliography

1. J.ALLEN : Natural language understanding, Benjamin/Cummings Publisher, 2nd ed., 1995.
2. E. CHARNIAK: Statistical language learning, MIT press, 1996.
3. B.CARPENTER: ALE: The attribute logic engine. User's guide. Carnegie Mellon University,1994.
4. D.FEHRER et al: Description logics for natural language processing. In Proc. of the 1994 Description Logic Workshop (DL'94), 1994.
5. H. HELBIG: Knowledge Representation and the Semantics of Natural Language, Springer, 2006.
6. D.JURAFSKY, J.MARTIN: Speech and language processing, Prentice Hall, 2000.
7. C.MANNING, H.SCHUTZE: Foundation of statistical natural language processing, MIT, 1999.
8. R. MITKOV(ed): The Oxford Handbook of Computational Linguistics, Oxford University Press, 2003.
9. D.TATAR: Inteligenta artificiala: demonstrare automata de teoreme, prelucrarea limbajului natural, Editura Albastra, Microinformatica, 2001.
10. D. TATAR: Inteligenta artificiala. Aplicatii in prelucrarea limbajului natural, Editura Albastra, Microinformatica, 2003, ISBN 973-650-100-01.

8.2 Seminar / laboratory	Teaching methods	Remarks
1. Work with WordNet, Romanian WordNet and WordNetSimilarity. Work with dedicated parsers and taggers (Stanford, CST tools, Racai tools)	Documentation on electronic platforms, explanation, dialogue, case studies	The seminar/lab is structured as 2 hours classes every second week
2. Study of platforms and libraries from different programming languages that offer preprocessing functions for texts in Romanian and English	Documentation on electronic platforms, dialogue, case studies	

Work with dedicated tools for keyword extraction, summarization, anaphora resolution, sentiment analysis.		
3. Students' presentations of a NLP task/ tool.	Dialogue, debate	
4. Identify practical tasks in Romanian NLP. Choose the NLP task, study different approaches, choose the approach that will be implemented. Search for the input data specific to the chosen task. Develop resources for Romanian NLP tasks	Documentation on electronic platforms, dialogue, case studies	
5. Design and implementation of the NLP tool. Develop resources for Romanian NLP tasks	Explanation, dialogue, case studies	
6. Students' presentations of the practical projects.	Evaluation	
Bibliography 1. Rada Mihalcea: www.cs.unt.edu/~rada/downloads.html 2. Resurse lingvistice in limba romana: www.racai.ro		

9. Corroborating the content of the discipline with the expectations of the epistemic community, professional associations and representative employers within the field of the program

- The course respects the IEEE and ACM Curricula Recommendations for Computer Science studies;
- The course exists in the studying program of all major universities in Romania and abroad;
- The optimization of the search on Web, the interfaces in natural language and the recent aspects of text mining need a good understanding of Natural Language Processing.

10. Evaluation

Type of activity	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Share in the grade (%)
10.4 Course	- know to write an overview of a specific NLP task	Theoretical report – presentation of a NLP task.	35%
10.5 Seminar/lab activities	- be able to implement course algorithms	Practical project - implementation of a NLP tool.	35%
	- be able to apply theoretical concepts in practical tasks	Develop resources for Romanian NLP tasks	20%
10.6 Activity	- activity during courses and labs	Active attendance	10%
10.7 Minimum performance standards			
➢ The final grade to be at least 5 (from a scale of 1 to 10).			

Date	Signature of course coordinator	Signature of seminar coordinator
27.04.2023	Lect. Ph.D. Lupea Mihaiela	Lect. Ph.D. Lupea Mihaiela

Date of approval	Signature of the head of department
.....	Prof. Ph.D. Dioşan Laura