

## SYLLABUS

### 1. Information regarding the programme

1.1 Higher education institution	<b>Babeş-Bolyai University</b>
1.2 Faculty	<b>Faculty of Mathematics and Computer Science</b>
1.3 Department	<b>Department of Computer Science</b>
1.4 Field of study	<b>Computer Science</b>
1.5 Study cycle	<b>Master</b>
1.6 Study programme / Qualification	<b>Computer Science/ High Performance Computing</b>

### 2. Information regarding the discipline

2.1 Name of the discipline	<b>Knowledge Based Systems and Language Technology</b>						
2.2 Course coordinator	<b>Lecturer Ph.D. Lupea Mihaiela</b>						
2.3 Seminar coordinator	<b>Lecturer Ph.D. Lupea Mihaiela</b>						
2.4. Year of study	<b>1</b>	2.5 Semester	<b>2</b>	2.6. Type of evaluation	<b>exam</b>	2.7 Type of discipline	<b>compulsory</b>

### 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 sem +1pr
3.4 Total hours in the curriculum	56	Of which: 3.5 course	28	3.6 seminar/laboratory	28
Time allotment:					hours
Learning using manual, course support, bibliography, course notes					25
Additional documentation (in libraries, on electronic platforms, field documentation)					15
Preparation for seminars/labs, homework, papers, portfolios and essays					25
Tutorship					10
Evaluations					15
Other activities: individual project					29
3.7 Total individual study hours	119				
3.8 Total hours per semester	175				
3.9 Number of ECTS credits	7				

### 4. Prerequisites (if necessary)

4.1. curriculum	Formal languages, Data structures, Machine learning
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	<ul style="list-style-type: none"> <li>Laboratory with computers; high level programming language environment (.NET or any Java environment a.s.o.)</li> </ul>

## 6. Specific competencies acquired

<b>Professional competencies</b>	<ul style="list-style-type: none"> <li>• Assimilation of mathematical concepts and formal models to understand, verify and validate software systems ;</li> <li>• Advanced ability to approach, model and solve phenomena and problems from natural language and economy using fundamental knowledge from mathematics and computer science;</li> <li>• Ability to approach and solve complex problems using various techniques of computational intelligence;</li> <li>• Proficient use of methodologies and tools specific to programming languages and software systems.</li> </ul>
<b>Transversal competencies</b>	<ul style="list-style-type: none"> <li>• Eitic and fair behavior, commitment to professional deontology</li> <li>• Team work capabilities; able to fulfill different roles</li> <li>• Professional communication skills; concise and precise description, both oral and written, of professional results , negotiation abilities;</li> <li>• Antepreneurial skills; working with economical knowledge; continuous learning</li> <li>• Good English communication skills</li> </ul>

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

7.1 General objective of the discipline	<ul style="list-style-type: none"> <li>• To introduce the basic principles, technologies and applications of Language Technology (LT) and Knowledge based systems.</li> <li>• To understand the current state of the art in LT in order to realize original research in LT.</li> </ul>
7.2 Specific objective of the discipline	<ul style="list-style-type: none"> <li>• Apply and use formal models (logics, grammars, parsing), statistic models (HMM), artificial intelligence algorithms (clustering, machine learning) and techniques (unsupervised, supervised) to solve different tasks at the syntactic level (POS-tagging, parsing, chunking), and semantic level (document summarization, word sense disambiguation, information extraction, anaphora resolution, sentiment analysis) in Natural Language Processing domain.</li> </ul>

## 8. Content

8.1 Course	Teaching methods	Remarks
<b>Course 1.</b> Natural Language Processing (NLP): stages, domains, applications.	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 2.</b> - Part-of-speech tagging. - WordNet: knowledge structure, semantic relations, lexical relations, applications, interfaces; corpora.	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 3. Text representation and text classification</b>	Exposure: description, explanation, examples, debate, dialogue	

<b>Course 4. Syntactic parsing</b> - grammar rules for English and sentence level construction; - Cocke-Kasami-Yonger (CKY) algorithm; - Probabilistic Context-Free Grammars (PCFG); - Probabilistic CKY algorithm	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 5. Hidden Markov Model (HMM)</b> - application to part-of-speech tagging	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 6. Keywords extraction</b>	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 7. Document summarization</b>	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 8. Opinion mining/Sentiment analysis</b>	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 9. Students' presentations of the research reports</b>	Debate, dialog	
<b>Course 10. Anaphora and co-reference resolution</b>	Exposure: description, explanation, examples, debate, dialogue Debate, dialog	
<b>Course 11. Word Sense Disambiguation</b>	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 12. Textual entailment</b>	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 13. Information extraction</b>	Exposure: description, explanation, examples, debate, dialogue	
<b>Course 14. Students' presentations of the practical project</b>	Debate, dialog	

## 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 tool. 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, explanation, dialogue, case studies	
3. Work with dedicated tools for information summarization, anaphora and co-reference resolution, sentiment analysis. Identify practical NLP tasks in Romanian language	Documentation on electronic platforms, explanation, dialogue, case studies	
4. Students' presentations of the research reports	Dialogue, debate	
5. Discussions about the practical projects	Explanation, dialogue, case studies	
6. Develop resources for Romanian NLP tasks	Documentation on electronic platforms, dialogue, case studies	
7. Students' presentations of the practical projects	Evaluation	
<b>Bibliography</b> 1. Rada Mihalcea: <a href="http://www.cs.unt.edu/~rada/downloads.html">www.cs.unt.edu/~rada/downloads.html</a> 2. Resurse lingvistice in limba romana: <a href="http://www.racai.ro">www.racai.ro</a>		

**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**

<ul style="list-style-type: none"> <li>• The course respects the IEEE and ACM Curricula Recommendations for Computer Science studies;</li> <li>• The course exists in the studying program of all major universities in Romania and abroad;</li> <li>• 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.</li> </ul>
---

**10. Evaluation**

Type of activity	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Share in the grade (%)
10.4 Course	- know the theoretical concepts of the domain; - apply the course methods, algorithms in problem solving	Written exam	20%
10.5 Seminar/lab activities	- know to synthesize and compare different approaches/results of the same studied subject.	Theoretical paper based on recent research papers in NLP domain;	35%
	- be able to implement course algorithms	Practical project - implementation of a NLP tool based on the studied methods	35%
	- be able to apply theoretical concepts in practical tasks	Develop resources for Romanian NLP tasks.	10%
10.6 Minimum performance standards			
➤ The final grade to be at least 5 (from a scale of 1 to 10) .			

Date

27.04.2023

Signature of course coordinator

Lect. Ph.D. Lupea Mihaiela

Signature of seminar coordinator

Lect. Ph.D. Lupea Mihaiela

Date of approval

.....

Signature of the head of department

Prof. Ph.D. Dioşan Laura