

SYLLABUS

Knowledge Based Systems and Language Technology

University year 2025-2026

1. Information regarding the programme

1.1. Higher education institution	Babeş-Bolyai University Cluj-Napoca
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	Master
1.6. Study programme/Qualification	Computer Science/ Applied Computational Intelligence
1.7. Form of education	Full time

2. Information regarding the discipline

2.1. Name of the discipline		Knowledge Based Systems and Language Technology					Discipline code		MME8044		
2.2. Course coordinator					Lecturer Ph.D. Lupea Mihaela						
2.3. Seminar coordinator					Lecturer Ph.D. Lupea Mihaela						
2.4. Year of study		1	2.5. Semester		2	2.6. Type of evaluation		E	2.7. Discipline regime		compulsory

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/project	1sem +1pr
3.4. Total hours in the curriculum	56	of which: 3.5 course	28	3.6 seminar/laboratory/project	28
Time allotment for individual study (ID) and self-study activities (SA)					hours
Learning using manual, course support, bibliography, course notes (SA)					30
Additional documentation (in libraries, on electronic platforms, field documentation)					20
Preparation for seminars/labs, homework, papers, portfolios and essays					30
Tutorship					10
Evaluations					20
Other activities: individual project					34
3.7. Total individual study hours	144				
3.8. Total hours per semester	200				
3.9. Number of ECTS credits	8				

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	Laboratory with computers; high level programming language environment (.NET or any Java environment a.s.o.)

6.1. Specific competencies acquired ¹

Professional/essential competencies	• Assimilation of mathematical concepts and formal model/s to understand, verify and validate software systems; • Ability to approach and solve complex problems using various techniques of computational intelligence.
Transversal competencies	• Capability of information analysis and synthesis; • Etic and fair behaviour, commitment to professional deontology.

6.2. Learning outcomes

Knowledge	• The graduate has the necessary knowledge to devise, model and design of complex software applications in the field of computational intelligence • The graduate possesses the fundamental knowledge for modelling, being able to analyse real life problems and to translate them in concrete requirements and to design a corresponding software model • The graduate proves knowledge related to specifying the requirements of research activities in the domain of computer science in general and computational intelligence in particular and he/she understands the role of research in promoting progress
Skills	• The graduate has the ability of interdisciplinary vision between computer science subdomains in order to combine them in a software system computational intelligence • The graduate can apply advanced computational intelligence knowledge starting from a high level of abstraction and being able to offer implementation solutions for complex software system.
Responsibility and autonomy:	• The graduate proves abilities to work independently in order to obtain knowledge necessary for designing, managing and evaluating research activities in the field of computational intelligence • The graduate knows and respects the ethical and legal principles and rules in scientific research. • The graduate has the ability to combine information in different ways in order to form a positive attitude towards its his/her own development

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

7.1 General objective of the discipline	• To introduce the basic principles, technologies and applications of Language Technology (LT) and Knowledge based systems. • To understand the current state of the art in LT in order to realize original research in LT.
---	---

¹ One can choose either competences or learning outcomes, or both. If only one option is chosen, the row related to the other option will be deleted, and the kept one will be numbered 6.

7.2 Specific objective of the discipline	• 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 <i>syntactic level</i> (POS-tagging, parsing, chunking), and <i>semantic level</i> (document summarization, word sense disambiguation, information extraction, anaphora resolution, sentiment analysis) in Natural Language Processing domain. • Use LLMs (Large Language Models) to solve NLP tasks in Romanian and English languages.
---	---

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: knowledge structure, semantic relations, lexical relations, applications, interfaces; corpora.	Exposure: description, explanation, examples, debate, dialogue	
Course 3. Text representation and text classification	Exposure: description, explanation, examples, debate, dialogue	
Course 4. Syntactic parsing	Exposure: description, explanation, examples, debate, dialogue	
Course 5. Hidden Markov Model (HMM) -application to part-of-speech tagging	Exposure: description, explanation, examples, debate, dialogue	
Course 6. Keywords extraction	Exposure: description, explanation, examples, debate, dialogue	
Course 7. Document summarization	Exposure: description, explanation, examples, debate, dialogue	
Course 8. Opinion mining/Sentiment analysis	Exposure: description, explanation, examples, debate, dialogue	
Course 9. Students' presentations of the research reports	Debate, dialog, evaluation	
Course 10. Anaphora and co-reference resolution	Exposure: description, explanation, examples, debate, dialogue Debate, dialog	
Course 11. Word Sense Disambiguation	Exposure: description, explanation, examples, debate, dialogue	
Course 12. Textual entailment	Exposure: description, explanation, examples, debate, dialogue	
Course 13. Information extraction	Exposure: description, explanation, examples, debate, dialogue	
Course 14. Students' presentations of the practical project	Debate, dialog, evaluation	
Bibliography 1. J. ALLEN : Natural language understanding, Benjamin/Cummings Publisher, 2nd ed., 1995. 2. L. DENG, Y. LIU: Deep learning in Natural Language Processing, Springer Verlag, Singapore, 2018. 3. H. HELBIG: Knowledge Representation and the Semantics of Natural Language, Springer, 2006. 4. D. JURAFSKY, J. MARTIN: Speech and language processing, Prentice Hall, 2000. 5. C. MANNING, H. SCHUTZE: Foundation of statistical natural language processing, MIT, 1999. 6. R. MITKOV(ed): The Oxford Handbook of Computational Linguistics, Oxford University Press, 2003. 7. G. PAAS, S. Giesselbach: Foundation Models for Natural Language Processing. Pre-trained Language Models Integrating Media, Springer, Berlin, 2022.		

8. D. ROTHMAN: Transformers for Natural Language Processing, Build, train, and fine-tune deep neural network architecture for NLP with Python, PyTorch, TensorFlow, BERT, and GPT-3, Second edition, Packt Publishing, 2022.		
9. D. TATAR: Inteligenta artificiala. Aplicatii in prelucrarea limbajului natural, Editura Albastra, Microinformatica, 2003.		
10. S. VAJJALA, B. MAJUMDER, A. GUPTA, H. SURANA: Practical Natural Language Processing. A comprehensive Guide to Building Real-World NLP Systems, O'REILLY. 2020.		
8.2 Seminar / laboratory	Teaching methods	Remarks
1. Study of platforms and libraries from different programming languages that offer preprocessing functions for texts in Romanian and English languages.	Documentation on electronic platforms, explanation, dialogue, case studies	The seminar/lab is structured as 2 hours classes every second week.
2. Study of the LLMs (large language models) for English and Romanian languages. Work with WordNet, Romanian WordNet and WordNetSimilarity. Work with dedicated tools for keyword extraction, summarization, anaphora resolution, sentiment analysis.	Documentation on electronic platforms, explanation, dialogue, case studies	
3. Identify practical tasks in Romanian NLP. Choose the NLP task to be solved, study different approaches, choose the approach that will be implemented. Search for the input data specific to the chosen task.	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, case studies.	
7. 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

Activity type	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Percentage of final grade
10.4 Course	- Know the theoretical concepts of the domain; - Apply the course methods, algorithms in problem solving. - Know to synthesize and compare different approaches of the same studied subject.	Written exam <i>or</i> theoretical report based on recent research papers in NLP domain.	40%
10.5 Seminar/laboratory	Be able to implement course algorithms	Practical project - implementation of an NLP tool based on the studied methods	40%
	Be able to apply theoretical concepts in practical tasks	Develop resources for Romanian NLP tasks.	20%
10.6 Minimum standard of performance			

- The final grade to be at least 5 (from a scale of 1 to 10).

11. Labels ODD (Sustainable Development Goals)²

Not applicable.

Date:
15.04.2025

Signature of course coordinator

Lect. Ph.D. Lupea Mihaela

Signature of seminar coordinator

Lect. Ph.D. Lupea Mihaela

Date of approval:
...

Signature of the head of department

Assoc.Prof.PhD. Adrian STERCA

² Keep only the labels that, according to the [Procedure for applying ODD labels in the academic process](#), suit the discipline and delete the others, including the general one for *Sustainable Development* – if not applicable. If no label describes the discipline, delete them all and write „*Not applicable.*”.