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	Master
1.6 Study programme / Qualification	Applied Computational Intelligence

2. Information regarding the discipline

2.1 Name of the discipline	Applications of Computational Linguistics						
2.2 Course coordinator	Lecturer PhD. Dana Lupsa						
2.3 Seminar coordinator	Lecturer PhD. Dana Lupsa						
2.4. Year of study	**2**	2.5 Semester	**3**	2.6. Type of evaluation	**E**	2.7 Type of discipline	**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	1sem +1 pr
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					30
Additional documentation (in libraries, on electronic platforms, field documentation)					44
Preparation for seminars/labs, homework, papers, portfolios and essays					30
Tutorship					20
Evaluations					20
Other activities:					

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	
4.2. competencies	• Average programming skills • Knowledge of data structures

5. Conditions (if necessary)

5.1. for the course	•
5.2. for the seminar /lab activities	•

6. Specific competencies acquired

Professional competencies	• Knowledge, understanding and use of IR concepts and their algorithms • Knowledge, understanding and use of of MT systems
Transversal competencies	• Perform Internet-based research. • Ability to use techniques specific to information retrieval and machine translation.

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

7.1 General objective of the discipline	• Understand how IR and MT systems works • Identify techniques for information retrieval, language translation
7.2 Specific objective of the discipline	• Explain basic information storage and retrieval concepts. • Describe what issues are specific to efficient information retrieval. • Perform Internet-based research. • Design and implement a small to medium size information storage and retrieval system, or digital library. • Understand techniques for information retrieval, language translation • Study the usage of corpora. Identify examples of corpora for MT

8. Content

8.1 Course	Teaching methods	Remarks
1. Ontologies	Exposure, examples, discussion	
2. Information retrieval. A first view. Vector space model (VSM)	Exposure, examples, discussion	
3. Boolean model. Extensions.	Exposure, examples, discussion	
4. Information Storage and Retrieval: Inverted Index; problems related to MultiWord expressions	Exposure, examples, discussion	
5. 5.1 Statistical properties of texts. 5.2 Vector similarity: issues	Exposure, description, explanation, examples, case studies	
6. 6.1 Dictionaries and tolerant retrieval 6.2 Meta-data, fields and zones	Exposure, description, explanation, examples, case studies	
7. Evaluation in IR	Exposure, description, explanation,	

	examples, case studies	
8. Semantic aspects	Exposure, description, explanation, examples, case studies	
9. Relevance feedback. Query expansion.	Exposure, description, explanation, examples, case studies	
10. Probabilistic retrieval	Exposure, examples, discussion	
11. Language models	Exposure, examples, discussion	
12. Web as graph: link analysis		
13. Efficient scoring and ranking: issues. Inexact top k retrieval	Description, examples, case studies, discussion	
14. Machine translation (MT).	Exposure, description, explanation, examples, case studies, discussion	

Bibliography

1. ALLEN, J.F. *Natural Language Semantics*, Wiley-Blackwell; 1 edition, 2001
2. D. Arnold, L. Balkan, S. Meijer, R. Humphreys, L. Sadler, *Machine Translation: An Introductory Guide, Manchester*, UK: NEC Blackwell, 1994.
3. R. BAEZA-YATES, B. RIBEIRO-NETO , *Modern Information Retrieval*, Addison-Wesley, 1999
4. E. CHARNIAK: *Statistical language learning*, MIT Press, 1996.
5. O. DAMERON, *Ontology-based methods for analyzing life science data*. Bioinformatics, Univ. Rennes , 2016
6. C.MANNING, H.SCHUTZE, *Foundation of statistical natural language processing*, MIT, 1999.
7. C. MANNING, P. RAGHAVAN, H. SCHUTZE, *Introduction to Information Retrieval*, Cambridge University Press, 2008.
8. R. MITKOV ed., *The Oxford Handbook of Computational Linguistics* (Oxford Handbooks in Linguistics), 2005

8.2 Seminar / laboratory	Teaching methods	Remarks
1,2: Knowledge representation in ontologies. Examples	Dialogue, examples	
3: Information in text – experiments: retrieval and ranking examples	Dialogue, examples	
4,5: Applications of theoretical techniques. Experiments.	Dialogue, case studies, examples	
6,7: Recent research in CL	Discussion, examples	

Bibliography

1. R. Mitkov (Ed), Oxford Handbook of Computational Linguistics. Oxford University Press, 2003.
2. C.D. Manning, P. Raghavan, H. Schütze,. Introduction to Information Retrieval. Cambridge, England: Cambridge University Press, 2008.

http://nlp.stanford.edu/IR-book/html/htmledition/irbook.html
3. http://www.mt-archive.info/
4. http://www.statmt.org/
5. web site: https://protege.stanford.edu/

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

- Two fields of CL with immediate applications in real life are IR and MT. IR systems are used on an everyday basis by a wide variety of users. The Internet has proven to be a huge stimulus for MT, with hundreds of millions of pages of text and an increasingly global -- and linguistically diverse – public.
- The course respects ACM Curriculla Recommendations for Computer Science studies

10. Evaluation

Type of activity	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Share in the grade (%)
10.4 Course	- know the basic principle of the domain; - apply the studied concepts	Written exam	40%
10.5 Seminar/lab activities	- apply the studied concepts - make experiments and solve problem	Oral presentation / interaction	10%
		Research report (presentations and experiments)	50 %
		There can be also other activities evaluated as bonus points	
10.7 Minimum performance standards			
➢ At least grade 5 (from a scale of 1 to 10) at both written exam and final computed grade			

Date	Signature of course coordinator	Signature of seminar coordinator
.....	lecturer PhD Dana Lupsa	lecturer PhD Dana Lupsa

Date of approval ..... Signature of the head of department .....