

LINGVISTICA COMPUTAȚIONALĂ

1. Date despre unitatea de curs/modul

| | | | | | |
|-------------------------------|---|------------------------|-------------------------------------|-----------------------------------|---------------------|
| Facultatea | Calculatoare, Informatică și Microelectronică | | | | |
| Catedra/departamentul | Informatică și Ingineria Sistemelor | | | | |
| Ciclul de studii | Studii superioare de licență, ciclul I | | | | |
| Programul de studiu | 0613.5 Informatică Aplicată | | | | |
| Anul de studiu | Semestrul | Tip de evaluare | Categoria formativă | Categoria de opționalitate | Credite ECTS |
| II (învățământ cu frecvență); | 4 | E | S – unitate de curs de specialitate | O - unitate de curs obligatorie | 4 |

2. Timpul total estimat

| | | | | | |
|-----------------------------------|-----------------|-------------------|-------------------|-------------------------------|---------------------|
| Total ore în planul de învățământ | Din care | | | | |
| | Ore auditoriale | | Lucrul individual | | |
| | Curs | Laborator/seminar | Proiect de an | Studiul materialului teoretic | Pregătire aplicații |
| 120 | 30 | 30 | | 30 | 30 |

3. Precondiții de acces la unitatea de curs/modul

| | |
|--------------------------------|---|
| Conform planului de învățământ | Tehnologii ale limbajului natural, Lingvistica teoretică. |
| Conform competențelor | Posedarea teoriei generale a limbajului. |

4. Condiții de desfășurare a procesului educațional pentru

| | |
|-----------|---|
| Curs | Pentru prezentarea materialului teoretic în sala de curs este nevoie de proiector și calculator. Nu vor fi tolerate întârzierile studenților, precum și convorbirile telefonice în timpul cursului. |
| Laborator | Studenții vor perfecta rapoarte conform condițiilor impuse de indicațiile metodice. Termenul de predare a lucrării de laborator – o săptămână după finalizarea acesteia. Pentru predarea cu întârziere a lucrării aceasta se depunceață cu 1pct./săptămână de întârziere. |

5. Competențe specifice acumulate

| | |
|-------------------------|---|
| Competențe profesionale | CPL1. Proiectarea aplicațiilor K1 Tehnici de modelare a cerințelor și tehnici de analiză a nevoilor. K3 Metricile care se referă la dezvoltarea aplicațiilor. |
| | CPL 2. Proiectarea și dezvoltarea aplicațiilor K1 Programe /module software adecvate. K3 Proiectarea funcțională și tehnică. K4 Tehnologiile de ultimă oră. |
| | CPL3. Integrarea componentelor K3 Tehnici de interfațare între module, sisteme și componente. K5 Instrumente de dezvoltare (ex. mediul de dezvoltare, gestionare, control al modificărilor și accesul la codul sursă). |
| Competențe transversale | CTL 1. Autonomie și responsabilitate. Demonstrează executarea responsabilă a sarcinilor profesionale în condiții de autonomie. |

| | |
|--|---|
| | <p>CTL 2. Interacțiune socială. Execută rolurile și activitățile specifice muncii în echipă și distribuie sarcinile între membri pe niveluri subordonate.</p> <p>CTL 3. Dezvoltare personală și profesională. Conștientizează nevoia de formare continuă cu utilizarea eficientă a resurselor și tehnicilor de învățare pentru dezvoltarea personală și profesională.</p> |
|--|---|

6. Obiectivele unității de curs/modulului

| | |
|-----------------------|--|
| Obiectivul general | Pregătirea și calificarea tinerilor specialiști în lingvistica computațională. |
| Obiectivele specifice | <p>să posede în mare măsură teoria generală a limbii;</p> <p>să fie capabili să soluționeze diverse probleme legate de aplicarea cunoștințelor din domeniul modelării limbajului și a proceselor de comunicare verbală;</p> <p>să poată utiliza în cadrul activității sale aparatul logico-formal matematic;</p> <p>să învețe și să utilizeze instrumentele existente de procesare a limbajului natural.</p> <p>familiarizarea cu noile tehnologii și programe lingvistice bazate pe computer.</p> |

7. Conținutul unității de curs/modulului

| Tematica activităților didactice | Numărul de ore | |
|---|-------------------------------|--|
| | învățământ cu frecvență | |
| Tematica prelegerilor | | |
| T1. Introducere în LC. Conceptul de LC și componentele sale. | 2 | |
| T2. Semantica structurală. Limbaj-obiect – metalimbaj. Sememul și semul. Structura sememului: semantem, clasem, virtuem. Coerența semantică. Izotopia. | 4 | |
| T3. Semantica transformațională. Lexiconul și regulile. Formative: mărci semantice și mărci distinctive. Restricții selective. Descrierea sensului unui formativ. Lecțiunile și amalgamarea lor. Clasificarea enunțurilor în semantica transformațională. | 4 | |
| T4. Lingvistica corpusului. Noțiuni de corpus. Tipuri de corpus. Necesitatea lucrului pe corpus. Elaborarea unui corpus. Segmentarea textelor. Adnotarea corpusului. Tipuri de adnotare a corpusului. | 4 | |
| T5. Adnotarea morfosintactică a corpusului. POS tagging. Seturi de etichete. Atribute și valori ale PDV. | 4 | |
| T6. Adnotarea sintactică a corpusului. Dependency Grammar Annotator. Relații de dependență în limba română. | 4 | |
| T7. Adnotarea semantică a corpusului. Lexiconul. Mărcile semantice. Ambiguitatea. | 4 | |
| T8. Dezambiguizarea sensurilor cuvintelor (WSD): noțiuni, principii. Abordarea bazată pe dicționar. Combinarea surselor de cunoștințe. Dezambiguizarea sensurilor cuvintelor (WSD): abordarea statistică și cea machine learning. Aplicații ale WSD. | 2 | |
| T9. Noțiuni de ontologie. Tipuri de ontologii. Utilizarea ontologiilor. Instrumente de creare a unei ontologii. Elaborarea unei ontologii cu ajutorul editorului de creare a sistemelor inteligente Protégé. | 2 | |
| Total prelegeri: | 30 | |
| Tematica activităților didactice | Numărul de ore | |

| | învățământ cu frecvență | |
|---|-------------------------------|-----------|
| Tematica lucrărilor de laborator/seminarelor | | |
| LL1. Analiza funcțională, sintagma-tică și transformațională a textului. | 8 | |
| LL2. Adnotarea morfosintactică a textului în baza setului de etichete elaborat. | 6 | |
| LL3. Adnotarea sintactică și semantică a textului în baza gramaticilor de dependență. | 8 | |
| LL4. Elaborarea unei ontologii în Protégé. | 8 | |
| Total lucrări de laborator/seminare: | | 30 |

8. Referințe bibliografice

| | |
|--------------|---|
| Principale | <ol style="list-style-type: none"> 1. Jurafsky, D., J. H. Martin. <i>Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition</i>. Prentice-Hall, 2000; www.cs.colorado.edu/~martin/slp.html . 2. Mitkov Ruslan. The Oxford Handbook on Computational Linguistics. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199276349.001.0001/oxfordhb-9780199276349 3. <i>Limba română în societatea informațională</i>. Societatea cunoașterii. Coordonatori: Dan Tufiș, Florin Gh. Filip. Academia Română, Institutul de cercetări pentru Inteligență Artificială. Ed. Expert, 2002. www.racai.ro 4. Igor A. Bolshakov and Alexander Gelbukh. Computational Linguistics. Models, Resources, Applications. In www.Gelbukh.com/clbook. 5. Saussure, Ferdinand de. <i>Course de linguistique generale</i>. 6. Chomsky, N. <i>Syntactic Structures</i>. The Hague: Mouton, 1957. 7. Ionescu, Emil. Manual de lingvistică generală. Editura ALL, București 1997. 8. An introduction to NLP. http://www.cs.bham.ac.uk/~pxc/nlpa/2002/AI-HO-IntroNLP.html 9. Allen, J. <i>Natural Language Understanding</i>. The Benjamin / Cummings Publ., Amsterdam, Bonn, Sidney, Singapore, Tokyo, Madrid, 1995. 10. Steele, J. (ed). <i>Meaning–Text Theory. Linguistics, Lexicography, and Implications</i>. University of Ottawa Press, Ottawa, 1990. 11. Proiectul EAGLES: http://www.eagles.org. 12. Cristea, D.; Ide, N.; Marcu, D.; Tablan, V. (1999): Discourse Structure and Co-Reference: An Empirical Study. În <i>Proceedings of the Workshop on the Relation Between Discourse Structure and Reference, ACL '99</i>, University of Maryland, 8 pag. www.info.uaic.ro/cristea.html. 13. Ruslan Mitkov. Anafora resolution. http://www.wlv.ac.uk/clg/mitkov.html 14. BUSUIOC, Ileana; CUCU, Mădălina. Introducere în terminologie. http://www.unibuc.ro/eBooks/filologie/terminologie/frameset.htm 15. Fillmore, C. J. <i>The case for case</i>. In: Bach, E., and B.T. Halm (eds.), <i>Universals in linguistic theory</i>. NY, Chicago, San Francisco, 1968. |
| Suplimentare | <ol style="list-style-type: none"> 1. http://protege.stanford.edu/ 2. http://bringiton.info.uaic.ro/proiecte/proiect24 3. https://ro.wikipedia.org/wiki/Sisteme_de_tip_%C3%AEntrebare_-_r%C4%83spuns 4. http://anale.feaa.uaic.ro/anale/resurse/41_Andone I - Ontologiile si modelarea informationala a intreprinderii.pdf 5. http://thor.info.uaic.ro/~busaco/publications/ontologii.html 6. http://turing.cs.pub.ro/webs_07/Prezentari Master/paperManagOnto.doc 7. http://revistaie.ase.ro/content/23/Niculescu.pdf 8. http://profs.info.uaic.ro/~alaiba/mw/index.php?title=Concepte_%C3%AEn_procesarea |

| | |
|--|---|
| | <p>limbaj_natural</p> <ol style="list-style-type: none"> 9. http://www.phobos.ro/roric/Ro/nlp.html 10. https://www.cs.cmu.edu/~mihaib/articole/ai/ai-html.html 11. https://en.wikipedia.org/wiki/Machine_learning 12. http://stats.stackexchange.com/questions/5026/what-is-the-difference-between-data-mining-statistics-machine-learning-and-ai 13. https://ro.wikipedia.org/wiki/Sumarizare_automat%C4%83 14. http://www.cs.cmu.edu/~nasmith/LS2/das-martins.07.pdf 15. http://www.isi.edu/natural-language/projects/SUMMARIST.html 16. https://crlrc.wordpress.com/about/descriere/ 17. http://www.phobos.ro/roric/Ro/DGA/Doc/annotation.htm 18. http://www.cnaa.md/thesis/6478/ 19. http://andrei.clubcisco.ro/cursuri/f/f-sym/4ioc/Prelucrarea%20documentelor%20folosind%20XML.pdf 20. http://consilr.info.uaic.ro/editii%20anterioare/2010/prezentari/14_CenelAugustoPerez.pdf 21. http://www.racai.ro/media/Tufis-Irimia-CONSILR2006.pdf 22. http://www.phobos.ro/roric/Ro/DGA/dga.html 23. http://www.cadredidactice.ub.ro/simonavarlan/files/2013/03/referat-2_iuliana-dobre1.pdf 24. http://www.phobos.ro/roric/Ro/DGA/Doc/index.htm 25. http://rochi.utcluj.ro/rrioc/articole/RRIOC-7-1-Florea.pdf 26. http://www.mcolhon.ro/didactics/courses/bc/Curs12.pdf 27. http://www.racai.ro/media/radu-ion-tezadoc.pdf 28. http://www.todaysoftmag.ro/article/82/semantic-web-scurta-introducere 29. http://www.scritub.com/stiinta/informatica/Rezolutia-anaforei14131783.php 30. http://www.unibuc.ro/prof/zafiu_r/docs/res/2011maranafora_bd_560791.pdf |
|--|---|

9. Evaluare

| Forma de învățământ | Periodică | | Curentă | Lucrul individual | Examen final |
|--|-------------|-------------|---------|-------------------|--------------|
| | Atestarea 1 | Atestarea 2 | | | |
| Cu frecvență | 15% | 15% | 15% | 15% | 40% |
| Standard minim de performanță | | | | | |
| <p>Prezența și activitatea la prelegeri și lucrări de laborator; Lucrare scrisă – test prin care se evaluează cunoștințele teoretice dobândite din tematica cursului și a laboratorului. Cerințe minime pentru nota 5: Cunoașterea elementelor fundamentale de teorie. Obținerea notei minime de „5” la fiecare din atestări și lucrări de laborator; Cerințe pentru nota 10: Cunoașterea tuturor elementelor de teorie predate la curs și la laborator. Realizarea tuturor temelor de laborator.</p> | | | | | |