F8-REG-CMCPU-02

1. **Date despre program**

# FIŞA DISCIPLINEI (IF)

# 2019-2020

**Lingvistică computațională**

|  |  |  |
| --- | --- | --- |
| 1.1 | Instituţia de învăţământ superior | **Universitatea din Piteşti** |
| 1.2 | Facultatea | **Teolgie, Litere, Istorie şi Arte** |
| 1.3 | Departamentul | **Limbi Straine Aplicate** |
| 1.4 | Domeniul de studii | **Filologie** |
| 1.5 | Ciclul de studii | **Master** |
| 1.6 | Programul de studiu / calificarea | ***Limbaje specializate şi traducere asistată de calculator*** |

1. **Date despre disciplină**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2.1 | Denumirea disciplinei | | | | | **Lingvistică computațională** | | | | | |
| 2.2 | Titularul activităţilor de curs | | | | | Conf. univ. dr. Ilinca Elena-Cristina | | | | | |
| 2.3 | Titularul activităţilor de seminar / laborator | | | | | - | | | | | |
| 2.4 | Anul de studii | I | 2.5 | Semestrul | II | 2.6 | Tipul de evaluare | E | 2.7 | Regimul disciplinei | O |

1. **Timpul total estimat**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3.1 | Număr de ore pe saptămână | 1 | 3.2 | din care curs | | 1 | 3.3 | S / L / P |  |
| 3.4 | Total ore din planul de înv. | 14 | 3.5 | din care curs | | 14 | 3.6 | S / L / P |  |
| **Distribuţia fondului de timp alocat studiului individual** | | | | | | | | | ore |
| Studiu după manual, suport de curs, bibliografie şi notiţe | | | | | | | | | 60 |
| Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate şi pe teren | | | | | | | | | 30 |
| Pregătire seminarii/laboratoare, teme, referate, portofolii, eseuri | | | | | | | | | 25 |
| Tutorat | | | | | | | | | 5 |
| Examinări | | | | | | | | | 2 |
| Alte activităţi ..... | | | | | | | | |  |
| 3.7 | Total ore studiu individual | | 122 | |  | | | | |
| **3.8** | **Total ore pe semestru** | | 136 | |
| **3.9** | **Număr de credite** | | 5 | |

1. **Precondiţii (acolo unde este cazul)**

|  |  |  |
| --- | --- | --- |
| 4.1 | De curriculum | Introducere în lingvistica corpusului.Exploatare avansată atehnologiilor informatice. |
| 4.2 | De competenţe | Competente dobândite la disciplinele Introducere în lingvistica corpusului. Exploatare avansată atehnologiilor informatice. |

1. **Condiţii (acolo unde este cazul)**

|  |  |  |
| --- | --- | --- |
| 5.1 | De desfăşurare a cursului |  |
| 5.2 | De desfăşurare a laboratorului | Sală cu minim 10 posturi de lucru dotate cu PC |

1. **Competenţe specifice acumulate**

|  |  |
| --- | --- |
| Competenţe profesionale | C1. Să efectueze activităţi de **documentare tematică** şi **terminologică ;**  C2. Să utilizeze la nivel avansat instrumente informatice de **prelucare automată a limbilor naturale**, de **gestiune de baze de date**, **traducere automată**, **traducere asistată de calculator**; |
| Competenţe transversale | **CT1.** Să dobândească abilități de cercetare interdisciplinară, să-şi dezvolte spiritul inovativ şi gândirea holistică; |

1. **Obiectivele disciplinei**

|  |  |
| --- | --- |
| 7.1 Obiectivul general al disciplinei | Obiectivul principal al disciplinei este pregătirea cursanților din punct de vedere practic în utilizarea și adaptarea uneltelor de prelucrarea a limbajului natural, precum și dezvoltarea capacității de a contribui la activitatea de cercetare și dezvoltare în acest domeniu. |
| 7.2 Obiectivele specifice | - prezentarea noțiunilor fundamentale ce țin de prelucrarea limbajului natural  - introducerea unor metode de statistică folosite în acest domeniu  - prezentarea unor unelte specifice acestui domeniu, folosite pe scară largă atât în cercetare cât și în producție  - însușirea unor elemente de bază privind sistemele de traducere automată |

1. **Conţinuturi**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **8.1. Curs** | | Nr. ore | Metode de predare | Observaţii Resurse folosite | |
| 1 | **Introducere în lingvistică computațională.** Prezentarea generală a domeniului. Noțiuni elementare de prelucrare a limbajului natural. Aplicații practice | 2 | Prelegere  Exemple  Descriere  Studiu de caz | Materiale didactice prezentate cu videoproiector  Posturi de lucru cu PC si acces la Internet | |
| 2 | **Noțiuni elementare de statistică.** Aplicarea modelelor statistice în prelucrarea limbajului natural. Noțiuni introductive de statistică | 2 |
| 3 | **Modele de limbă.** Calculul probabilitații pentru unigrame, bigrame și trigrame. Netezire statistică pentru evenimente necunoscute | 2 |
| 4 | **Etichetare morfosintactică**. Prezentare generală. Metode clasice folosite în etichetare morfosintactică: Modele Markov de ordinul întâi**.** | 2 |
| 5 | **Unelte pentru prelucrarea limbajului natural folosite în scop academic .** Prezentarea uneletelor de bază în prelucrarea limbajului natural. Stanford NLP Tools. Unelte RACAI pentru limba română | 2 |  |  | |
| 6 | **Sisteme de traducere automată.** Introducere. Sisteme bazate pe modele de suprafață – echivalenți de traducere. Descrierea sistemului MOSES. Evaluarea sistemelor de traducere automată: scorul BLEU și scorul METEOR | 2 |  |  | |
| 7 | **Extragerea automată a echivalenților de traducere.** Noțiunea de corpus paralel și corpus comparabil. Metode de aliniere automată | 2 |  |  | |
|  | | |  |  | |
| **Bibliografie**  R.H. Baayen. 2008. Analyzing Linguistic Data: A Practical Introduction to Statistics using R. Cambridge University Press.  Tufiș, D. 2009. Algorithms and Data Design Issues for Basic NLP Tools. *Language Engineering for Lesser-studied Languages*, 21, 3.  Boroș, T., Dumitrescu, S. D., Ion, R., Ștefănescu, D., and Tufiș, D. (2013). *Romanian-English statistical translation at RACAI*. ISSN 1843-911x, 81.  Boroş, T., Radu, I., & Tufiş, D. 2013. Large tagset labeling with Feed Forward Neural Networks. Case study on Romanian Language. In *Proceedings of ACL*.  Ilinca, C. 2019. *Lingvistică computațională* (note de curs), suport electronic  Paul, Michael J. 2012. Mixed membership Markov models for unsupervised conversation modeling. In Proceedings of the 2012 *Joint* *Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, EMNLP-CoNLL ’12, pages 94–104, Stroudsburg, PA.  Tran, Quan Hung, Ingrid Zukerman, and Gholamreza Haffari. 2016. Inter-document contextual language model. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 762–766.   [Joty](https://www.mitpressjournals.org/author/Joty%2C+Shafiq), Shafiq and Tasnim Mohiuddin. 2018. Modeling Speech Acts in Asynchronous Conversations: A Neural-CRF Approach, *Computational Linguistics*, Volume 44, Issue 4, December 2018 p.859-894  [Reiter](https://www.mitpressjournals.org/author/Reiter%2C+Ehud), Ehud. 2018. A Structured Review of the Validity of BLEU, *Computational Linguistics,* Volume 44, Issue 3, September 2018, p.393-401 | | | | | |

1. **Coroborarea conţinuturilor disciplinei cu aşteptările reprezentanţilor comunitaţii epistemice, asociaţiilor profesionale şi angajatori din domeniul aferent programului**

Competenţele dobândite la disciplină permit absolvenţilor să lucreze în domeniul aferent calificării Limbaje specializate şi traducere asistată de calculator

1. **Evaluare**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tip activitate | 10.1 Criterii de evaluare | | 10.2 Metode de evaluare | 10.3 Pondere din nota finală |
| 10.4 Curs | Evaluare finală | | Examen scris | 50% |
|  | Verificarea cunoștințelor | | Referat  Teme | 30%  20% |
| 10.6 Standard minim de performanţă | | - cunoașterea noțiunilor fundamentale privind prelucrarea limbajului natural  - cunoașterea uneltelor de bază și a pachetelor disponibile în scopuri academice  - abilitatea de a antrena și folosi un sistem de traducere automată | | |

Data completării Titular de curs, Titular de seminar / laborator,

20 septembrie 2019 Conf. univ. dr. Ilinca Elena-Cristina

Data aprobării în Consiliul departamentului Director de departament,

25 septembrie 2019 Conf. univ. dr. Cîţu Laura