

## Referate Computational Linguistics & NLP

-Liviu P Dinu-

Propuneri de teme:

1. Rezumare automata a textelor
2. POS-tagging
3. Parsare; **Procesare la nivel sintactic: gramatici de dependență; analiza sintactică de dependență; Stanford parser, etc**
4. **Resurse electronice (corpus, etc)**
5. Simplificare de text
6. Similaritate semantica
7. Semantica distributionala
8. Similaritate sintactica
9. Detectare Fake news
10. Deception detection
11. Hate speech
12. Detectarea ironiei
13. Analiza computationala a umorului
14. Authorship identification
15. Detectarea metaforei
16. Lingvistica istorica, abordari computationale: producere de cuvinte, detectare de cognates, proto-word, borrowing, etc
17. Language generation.
18. Social Media
19. Deep learning in NLP
20. Machine learning in NLP.
21. Traducere Automata (abordari, modele):
22. Dezambiguizare (state of the art, algoritmi, implementari)
23. Named Entity Recognition (abordari))
24. Collocation detection
25. Sentiment Analysis
26. Opinion Mining
27. Analiza optimism-pesimism
28. <https://naacl2018.wordpress.com/2018/01/14/test-of-time-paper-nominations-or-classic-computational-linguistics-papers/> (discutia unui articol din aceasta lista)
29. Topici recente de cercetare in domeniu (articole recente relevante din Computational Linguistics, ACL, COLING, EMNLP, NAACL, EACL, PNAS, etc) propunerile de best papers in ultimii 10 ani (Disponibile on-line la <http://aclweb.org/anthology-new/> )

30. Lista cu articolele de la LREC 2020 propuse pentru re-experimentare (<https://lrec2020.lrec-conf.org/en/reprolang2020/selected-tasks/>).

31. Aplicatii ale NLP

32. Alte prezentari legate de Computational Linguistics & NLP care nu se gasesc in aceasta lista, pe care sa le discutam in prealabil (trimis un mail in prealabil cu tema propusa).

### **Cerinte (recomandari ):**

Fiecare tema aleasa (prezentare) va urmari structura clasica a unui articol de cercetare:

-scurt rezumat (cateva randuri)

-ideea principala analizata

-related work: state of the art (SOTA) acolo unde este cazul, scurt istoric, rezultate recente

-In cazul prezentarii unui survey: metodologiile principale, explicate (selectie), fie in ordine cronologica, fie in ordinea SOTA, prezentarea avantajelor si dezavantajelor metodelor folosite

-studiu de caz (obligatoriu pentru echipele de 2-3)

-in cazul prezentarii unor aplicatii punctuale:descrierea metodei, compararea cu rezultatele din domeniu

-concluzii si future work, directii de imbunatatit

-bibliografie

Prezentarea se va face pe slide-uri (maxim 20-30). Va fi atasata o anexa cu explicatii ale prezentarii (1-2 pagini maxim).

Echipele sunt de 1, 2 sau maxim 3 colegi. Pentru echipele de 2-3 colegi obligatoriu si un studiu de caz (aplicatie, implementare, refacere de experimente);

Pentru orice clarificare, intrebare, nelamurire, etc, eu sunt on-line in cea mai mare parte a timpului, trimiteti email sau stabilim o discutie (hangout, skype, telefon, etc).

II. Fiecare tema poate fi aleasa de maxim 2-3 echipe.

III. **Propunere:** faceti voi o lista publica in care fiecare sa se inscrie cu ce tema isi alege?

Sau, sub fiecare tema de mai sus cum este aleasa de o echipa sa-si treaca numele (toti participantii) sub tema respectiva (poate e mai simplu asa).