

Tematica și bibliografia pentru concurs **Conferențiar 10**

- 1. Scurtă istorie a domeniului DH**
- 2. Tipuri de obiecte digitale și reprezentarea acestora standardizată (XML, TEI)**
- 3. Metode în DH:**
 - Colectarea și adnotarea de corpusuri
 - Trecerea textului în format digital – OCR-izare
 - Curățarea corpusurilor
 - Interogarea corpusurilor, expresii regulate, expresii multi-cuvânt
 - Procesarea automată a corpusurilor:
 - Tokenizare, stematizare și lematizare
 - Etichetare cu părți de vorbire
 - Cuvinte gramaticale
 - Analiza sintactică automată a textelor: parsare bazată pe constituenți, parsare de dependență
 - Analiza semantică a textelor:
 - formală (semantici Montague, lambda-calcul)
 - distribuțională (reprezentări vectoriale ale cuvintelor, propozițiilor și textelor, statice și contextuale, transformare și modele mari de limbaj)
 - Analiza propriu-zisă a corpusurilor:
 - Analiza cantitativă a corpusurilor
 - Analiza sentimentelor
 - Extragerea automată a temelor
 - Recunoașterea automată a entităților
 - Sarcini abordate prin tehnici de învățare automată tradițională și prin tehnici de învățare profundă
 - Evaluarea și interpretarea rezultatelor:
 - Metrice de evaluare a similarității între documente
 - Metrice de evaluare a învățării automate
 - Teste statistice
- 4. Instrumente și biblioteci în DH:**
 - Instrumente de analiză de textelor
 - Instrumente pentru vizualizarea datelor și a rețelelor
 - Biblioteci pentru procesare de text și vizualizare
 - LLM-uri disponibile ca instrumente
- 5. Aplicații în DH, precum:**
 - Analiza computațională a discursurilor politice ale președinților României

- Identificarea și clasificarea automată a postărilor care conțin învinovățirea victimelor abuzurilor sexuale
- Diferențe generaționale în exprimarea pe platforme de socializare
- Detectarea evoluției stilistice de la comunism la democrație în cazul operei publicistice a lui Solomon Marcus
- Măsurarea schimbării semantice a cuvintelor
- Identificarea automată a cuvintelor înrudite în limbile romanice
- Prezicerea automată a accentului în limba română
- Testarea abilităților creative ideationale și de limbaj ale modelelor de limbaj mari și compararea acestora cu abilitățile creative umane
- Testarea abilităților modelelor de limbaj mari de a pastșa opere literare și compararea acestora cu abilitățile umane
- Analiza contrastivă multimodală a *meme*-urilor de tip *brain rot* în italiană și română

Bibliografie:

1. Gardiner, E., & Musto, R. G. (2015). *The digital humanities: A primer for students and scholars*. Cambridge University Press.
2. Susan Schreibman, Ray Siemens, & John Unsworth (Eds.). (2016). *A new companion to digital humanities*. Wiley-Blackwell.
3. Dinu Anca Daniela. *Digitalizarea în științele umaniste. Metode și aplicații*. Editura Prouniversitaria, ISBN 978-606-26-2152-0, 2025.
4. Dinu, A., & Florescu, A. M. (2024). *An integrated benchmark for verbal creativity testing of LLMs and humans*. *Procedia Computer Science*, 246, 2902–2911.
5. Dinu, A., & Florescu, A. M. (2024b). “Comparing large language models verbal creativity to human verbal creativity”. In F. Dell’Orletta, A. Lenci, S. Montemagni, & R. Sprugnoli (Eds.), *Proceedings of the Tenth Italian Conference on Computational Linguistics (CLiC-it 2024)*. CEUR-WS.org.
6. Dinu, A., & Florescu, A.-M. (2025). “Testing language creativity of large language models and humans.” *Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities* (co-located with NAACL 2025), 426–436.
7. Dinu, A., & Niculescu, A. M. (2025). “Automatically identifying victim blaming on social media in cases of sexual assault, sexual harassment and abuse”. In *Proceedings of the 18th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA ’25)* (pp. 492–496). Association for Computing Machinery.
8. Dinu, A., & Teodorescu, L. (2025). “Generational differences in the use of Romanian language.” Presented at RADH 2025, Craiova, Romania.
9. Dinu, A., Dinu, L. P., & Dumitru, B. (2017). “On the stylistic evolution from communism to democracy: Solomon Marcus study case”. *Proceedings of RANLP 2017*, 201–207.

10. Dinu, A., Dobre, D. I., Moldovan, A.-C., & Nicolescu, D. (2022). "Computational analysis and author detection for political discourses of Romanian presidents". In M. Chitez, A. Dinu, L. Dinu, & M. Dobre (Eds.), *Recent advances in digital humanities. Romance language applications*. Peter Lang.
11. Dinu, A., Florescu, A. M., & Dinu, L. P. (2025a). "Analyzing large language models' pastiche ability: A case study on a 20th century Romanian author". *Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities*, 20–32.
12. Dinu, A., Florescu, A. M., & Dinu, L. P. (2025b). "Evaluating large language models' ability to pastiche literary texts." Presented at DTW: Digital Transformation Workshop, collocated with PETRA 2025.
13. Dinu, A., Florescu, A. M., & Tăbușcă, Ș. A. (2025c). "Uncovering the differences between LLM-generated and human-written answers to an ideational creativity test". *Proceedings of The 28th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2025)*.
14. Dinu, A., Florescu, A. M., Micluță-Câmpeanu, M., Tăbușcă, Ș. A., Creangă, C., & Mihail, A. (2025d). "Dissonant ballerinas and crafty carrots: A comparative multi-modal analysis of Italian brain rot". *Proceedings of CLiC-it 2025*, Cagliari, Italy.
15. Dinu, A., Florescu, A.-M., & Resceanu, A. (2025e). "A comparative approach to assessing linguistic creativity of large language models and humans". *Proceedings of The 28th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2025)*.
17. Dinu, L. P., Niculae, V., & Șulea, O.-M. (2012). "Pastiche detection based on stopword rankings: Exposing impersonators of a Romanian writer". *Proceedings of the Workshop on Computational Approaches to Deception Detection (EACL 2012)*, Avignon, France. Association for Computational Linguistics.
18. Dinu, L. P., Niculae, V., & Șulea, O.-M. (2013). "Romanian syllabication using machine learning". *Proceedings of the 16th International Conference on Text, Speech and Dialogue (TSD 2013)*, 450–456.
19. Dinu, L. P., Popescu, M., & Dinu, A. (2008). "Authorship identification of Romanian texts with controversial paternity". *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, 3392–3397. European Language Resources Association (ELRA).
20. Dinu, L. P., Uban, A. S., Cristea, A. M., Dinu, A., Iordache, I.-B., Georgescu, S., & Zoicas, L. (2023). "RoBoCoP: A comprehensive ROmance BOrrrowing COgnate package and benchmark for multilingual cognate identification". In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (pp. 7610–7629). Association for Computational Linguistics.