

Topics and Bibliography for **Associate Professor position 10 – promotion track**

- 1. A brief history of the Digital Humanities**
- 2. Types of digital objects and their standardized representation (XML, TEI)**
- 3. Methods in Digital Humanities:**
 - Corpus collection and annotation
 - Text digitization, Optical Character Representation
 - Corpus cleaning
 - Corpus querying: regular expressions, multi-word expressions
 - Automatic corpus processing:
 - Tokenization, stemming, and lemmatization
 - Part-of-speech tagging
 - Function words
 - Parsing: constituency parsing, dependency parsing
 - Semantic analysis of texts:
 - formal semantics (Montague semantics, lambda calculus)
 - distributional semantics (word embeddings, sentence embeddings; static and contextual embeddings; transformers and large language models)
 - Corpus analysis:
 - Quantitative analysis
 - Sentiment analysis
 - Topic modelling
 - Named entity recognition
 - Machine Learning and Deep Learning tasks
 - Results evaluation and interpretation
 - Document similarity metrics
 - Machine Learning evaluation metrics
 - Statistical tests
- 4. Tools and Libraries in Digital Humanities:**
 - Text analysis tools
 - Data and network visualization tools
 - Libraries for text processing and visualization
 - Large Language Models (LLMs) as tools
- 5. Digital Humanities applications, e.g.:**
 - A computational analysis of the political speeches of Romanian presidents
 - Automatically identifying victim blaming on social media in cases of sexual assault, sexual harassment and abuse

- Generational differences in language use on Romanian social media
- An automatic analysis of the evolution of Solomon Marcus's journalistic style from communism to democracy
- Measuring semantic change of lexical items
- Automatic identification of cognates in Romance languages
- Automatic prediction of stress in Romanian
- Testing the ideational and linguistic creativity abilities of large language models and comparing them with human abilities
- Testing the ability of large language models to produce literary pastiche and comparing it with human performance
- Multimodal contrastive analysis of “brain rot” memes in Italian and Romanian

Bibliography:

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*. [Digital Humanities: Methods and Applications] 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)*.
16. 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 BORrowing 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.