ACADEMIA ROMÂNĂ
SCOSAAR

*Anexa nr.3*

**AVIZAT,**

Director ŞCOALA DOCTORALĂ DE ŞTIINŢE MATEMATICE ŞI INFORMATICĂ

1. Îndeplinirea standardelor IOSUD superioare standardelor minimale naţionale* [] DA| [] | NU

2. Îndeplinirea standardelor IOSUD egale standardelor minimale naţionale* [] DA| [] | NU

# FIŞA DE ÎNDEPLINIRE A STANDARDELOR IOSUD

## FIŞA DE VERIFICARE
a îndeplinirii standardelor IOSUD

Candidat: BOROŞ TIBERIU

Data:16.12.2024

Semnătura:

*se va alege una dintre variante

**Autoevaluarea activității profesionale, de cercetare și a recunoașterii și impactului acesteia, conform Anexei nr. 1 la OMECTS 6129/2016**

**Perspectiva A – Etica cercetării**

Nu există evidențe din care să reiasă nerespectarea eticii cercetării prin atribuirea de rezultate, texte sau imagini copiate din alte surse.

**Perspectiva B – Producția științifică**

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boros, Radu Ion, Dan Tufiș | **Publicație**: ACL (2013) | **Clasă**: A* (CORE) | **Punctaj**: 12 |
| **Titlu**: Large tagset labeling using Feed Forward Neural Networks. Case study on Romanian Language | | | |
| **Referință:** BOROȘ, Tiberiu; ION, Radu; TUFIȘ, Dan. Large tagset labeling using feed forward neural networks. case study on romanian language. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2013. (pp. 692-700). | | | |
| **Link/DOI:** https://aclanthology.org/P13-1068.pdf | | | |

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boros | **Publicație**: RANLP (2013) | **Clasă**: C (CORE) | **Punctaj**: 2 |
| **Titlu**: A Unified Lexical Processing Framework Based on the Margin Infused Relaxed Algorithm. A Case Study on the Romanian Language | | | |
| **Referință:** BOROȘ, Tiberiu. A unified lexical processing framework based on the Margin Infused Relaxed Algorithm. A case study on the Romanian Language. In: *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*. 2013. (pp. 91-97). | | | |
| **Link/DOI:** https://aclanthology.org/R13-1012.pdf | | | |

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boroș, Adriana Stan, Oliver Watts, Stefan Dumitrescu | **Publicație**: LREC (2014) | **Clasă**: B (CORE) | **Punctaj**: 2 |
| **Titlu**: RSS-TOBI-A Prosodically Enhanced Romanian Speech Corpus | | | |
| **Referință:** BOROȘ, Tiberiu, et al. RSS-TOBI-a Prosodically Enhanced Romanian Speech Corpus. *Proceedings of 9th Language Resources and Evaluation Conference (LREC) 2014*, 2014, (pp. 316-320). | | | |
| **Link/DOI:** https://www.cstr.ed.ac.uk/downloads/publications/2014/boros-2014.pdf | | | |

| **Autori**: Tiberiu Boros, Sonia Pipa, Verginica Barbu Mititelu, Dan Tufis | **Publicație**: MWE Workshop (EACL) (2017) | **Clasă**: B (A) (CORE) | **Punctaj**: 1 |
|---|---|---|---|
| **Titlu**: A data-driven approach to verbal multiword expression detection. PARSEME Shared Task system description paper | | | |
| **Referință:** BOROȘ, Tiberiu, et al. A data-driven approach to verbal multiword expression detection. PARSEME Shared Task system description paper. In: *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*. 2017. (pp. 121-126) | | | |
| **Link/DOI:** https://aclanthology.org/W17-1716.pdf | | | |

| **Autori**: Tiberiu Boros, Radu Ion, Stefan Dumitrescu | **Publicație**: SSW (2012) (Interspeech) | **Clasă**: B (A) (CORE) | **Punctaj**: 2 |
|---|---|---|---|
| **Titlu**: The RACAI Text-to-Speech Synthesis System | | | |
| **Referință:** BOROȘ, Tiberiu; ION, Radu; DUMITRESCU, Ștefan Daniel. The RACAI text-to-speech synthesis system. *Blizzard Challenge - Speech Synthesis Workshop (SSW)*, 2013. (pp. 74-79). | | | |
| **Link/DOI:** http://dx.doi.org/10.21437/Blizzard.2013-12 | | | |

| **Autori**: Tiberiu Boroș, Stefan Dumitrescu, Adrian Zafiu, Dan Tufiș, Verginica Barbu Mititelu, Paul Ionut Văduva | **Publicație**: CONLL (2014) | **Clasă**: A (CORE) | **Punctaj**: 2 |
|---|---|---|---|
| **Titlu**: RACAI GEC–A hybrid approach to Grammatical Error Correction | | | |
| **Referință:** BOROȘ, Tiberiu, et al. RACAI GEC–a hybrid approach to grammatical error correction. In: *Proceedings of the Eighteenth Conference on Computational Natural Language Learning (CONLL)*. 2014. (pp. 43-48). | | | |
| **Link/DOI:** https://aclanthology.org/W14-1705.pdf | | | |

| **Autori**: Tiberiu Boros, Stefan Dumitrescu | **Publicație**: EANN (2013) | **Clasă**: C (CORE) | **Punctaj**: 2 |
|---|---|---|---|
| **Titlu**: Improving the RACAI Neural Network MSD Tagger | | | |
| **Referință:** BOROȘ, Tiberiu; DUMITRESCU, Stefan Daniel. Improving the racai neural network msd tagger. In: *Engineering Applications of Neural Networks: 14th International Conference, EANN 2013, Halkidiki, Greece, September 13-16, 2013 Proceedings, Part I 14*. Springer Berlin Heidelberg, 2013. (pp. 42-51). | | | |
| **Link/DOI:** https://www.researchgate.net/profile/Stefan-Dumitrescu-2/publication/290040964_Improving_the_RACAI_Neural_Network_MSD_Tagger/links/5b8bd0ce4585151fd142f643/Improving-the-RACAI-Neural-Network-MSD-Tagger.pdf | | | |

| **Autori**: Tiberiu Boros, Stefan Dumitrescu, Dan Tufis | **Publicație**: CONLL (2017) | **Clasă**: A (CORE) | **Punctaj**: 8 |
|---|---|---|---|
| **Titlu**: RACAI's Natural Language Processing pipeline for Universal Dependencies | | | |

| |
|---|
| **Referință:** DUMITRESCU, Ștefan Daniel; BOROȘ, Tiberiu; TUFIȘ, Dan. Racai's natural language processing pipeline for universal dependencies. In: *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. 2017. (pp. 174-181). |
| **Link/DOI:** http://dx.doi.org/10.18653/v1/K17-3018 |

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boros, Stefan Dumitrescu | **Publicație**: EANN (2017) | **Clasă**: C (CORE) | **Punctaj**: 2 |
| **Titlu**: A convolutional approach to multiword expression detection based on unsupervised distributed word representations and task-driven embedding of lexical features | | | |
| **Referință:** BOROS, Tiberiu; DUMITRESCU, Stefan Daniel. A convolutional approach to multiword expression detection based on unsupervised distributed word representations and task-driven embedding of lexical features. In: *Engineering Applications of Neural Networks: 18th International Conference, EANN 2017, Athens, Greece, August 25–27, 2017, Proceedings*. Springer International Publishing, 2017. (pp. 149-159). | | | |
| **Link/DOI:** http://dx.doi.org/10.1007/978-3-319-65172-9_13 | | | |

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boros, Stefan Dumitrescu, Sonia Pipa | **Publicație**: EACL (2017) | **Clasă**: A (CORE) | **Punctaj**: 8 |
| **Titlu**: CASSANDRA: A multipurpose configurable voice-enabled human-computer-interface | | | |
| **Referință:** BOROȘ, Tiberiu; DUMITRESCU, Ștefan Daniel; PIPA, Sonia. CASSANDRA: A multipurpose configurable voice-enabled human-computer-interface. In: *Proceedings of the Software Demonstrations of the 15th Conference of the European Chapter of the Association for Computational Linguistics*. 2017. (pp. 33-36). | | | |
| **Link/DOI:** https://aclanthology.org/E17-3009.pdf | | | |

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boros, Stefan Dumitrescu, Ruxandra Burtica | **Publicație**: CONLL (2018) | **Clasă**: A (CORE) | **Punctaj**: 8 |
| **Titlu**: NLP-Cube: End-to-end raw text processing with neural networks | | | |
| **Referință:** BOROȘ, Tiberiu; DUMITRESCU, Ștefan Daniel; BURTICA, Ruxandra. NLP-Cube: End-to-end raw text processing with neural networks. In: *Proceedings of the Conference of Natural Language Learning (CoNLL) 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. 2018. (pp. 171-179). | | | |
| **Link/DOI:** https://aclanthology.org/K18-2017.pdf | | | |

| **Autori**: Tiberiu Boros, Stefan Dumitrescu | **Publicație**: CONLL - SIGMORPHON (2017) | **Clasă**: A (CORE) | **Punctaj**: 8 |
|---|---|---|---|
| **Titlu**: Attention-free encoder decoder for morphological processing | | | |
| **Referință:** DUMITRESCU, Ştefan Daniel; BOROŞ, Tiberiu. Attention-free encoder decoder for morphological processing. In: *Proceedings of the CoNLL–SIGMORPHON 2018 Shared Task: Universal Morphological Reinflection*. 2018. (pp. 64-68). | | | |
| **Link/DOI:** http://dx.doi.org/10.18653/v1/K18-3007 | | | |

| **Autori**: Tiberiu Boros, Ruxandra Burtica | **Publicație**: LAW-MWE (COLING 2018) | **Clasă**: B (A) (CORE) | **Punctaj**: 2 |
|---|---|---|---|
| **Titlu**: GBD-NER at PARSEME shared task 2018: Multi-word expression detection using bidirectional long-short-term memory networks and graph-based decoding | | | |
| **Referință:** BOROŞ, Tiberiu; BURTICA, Ruxandra. GBD-NER at PARSEME shared task 2018: Multi-word expression detection using bidirectional long-short-term memory networks and graph-based decoding. In: *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*. 2018. (pp. 254-260). | | | |
| **Link/DOI:** https://aclanthology.org/W18-4928.pdf | | | |

| **Autori**: T Boros, A Cotaie, K Vikramjeet, V Malik, L Park, N Pachis | **Publicație**: IoTBDS 2021 | **Clasă**: C (CORE) | **Punctaj**: 0.5 |
|---|---|---|---|
| **Titlu**: A Principled Approach to Enriching Security-related Data for Running Processes through Statistics and Natural Language Processing | | | |
| **Referință:** BOROS, Tiberiu, et al. A Principled Approach to Enriching Security-related Data for Running Processes through Statistics and Natural Language Processing. In: *IoTBDS*. 2021. (pp. 140-147). | | | |
| **Link/DOI:** http://dx.doi.org/10.5220/0010381401400147 | | | |

| **Autori**: Tiberiu Boros, Andrei Cotaie | **Publicație**: IoTBDS 2022 | **Clasă**: C (CORE) | **Punctaj**: 2 |
|---|---|---|---|
| **Titlu**: Machine Learning and Feature Engineering for Detecting Living off the Land Attacks. | | | |
| **Referință:** BOROS, Tiberiu, et al. Machine Learning and Feature Engineering for Detecting Living off the Land Attacks. In: *IoTBDS*. 2022. (pp. 133-140). | | | |
| **Link/DOI:** http://dx.doi.org/10.5220/0011004500003194 | | | |

| **Autori**: Tiberiu Boros, Andrei Cotaie | **Publicație**: IoTBDS 2023 | **Clasă**: C (CORE) | **Punctaj**: 2 |
|---|---|---|---|
| **Titlu**: Deep Dive into Hunting for LotLs Using Machine Learning and Feature Engineering. | | | |

| |
|---|
| **Referință:** BOROS, Tiberiu; COTAIE, Andrei. Deep Dive into Hunting for LotLs Using Machine Learning and Feature Engineering. In: *IoTBDS*. 2023. (pp. 194-199). |
| **Link/DOI:** http://dx.doi.org/10.5220/0011968700003482 |

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boros, Stefan Dumitrescu, Ionut Mironica, Radu Chivereanu | **Publicatie**: Speech Synthesis Workshop 2023 (Interspeech Workshop) | **Clasa:** B (A) (CORE) | **Punctaj:** 1 |
| **Titlu**: Generative Adversarial Training for Text-to-Speech Synthesis Based on Raw Phonetic Input and Explicit Prosody Modelling | | | |
| **Referință:** BOROS, Tiberiu, et al. Generative Adversarial Training for Text-to-Speech Synthesis Based on Raw Phonetic Input and Explicit Prosody Modelling. In Proceedings 18th Blizzard Challenge Workshop - Speech Synthesis Workshop (SSW), 2023. (pp. 69-74). | | | |
| **Link/DOI:** http://dx.doi.org/10.21437/Blizzard.2023-9 | | | |

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boros, Marius Barbulescu | **Publicatie**: IoTBDS 2024 | **Clasa**: C (CORE) | **Punctaj**: 2 |
| **Titlu**: Hybrid Statistical Modeling for Anomaly Detection in Multi-Key Stores Based on Access Patterns | | | |
| **Referință:** BOROS, Tiberiu; BARBULESCU, Marius. Hybrid Statistical Modeling for Anomaly Detection in Multi-Key Stores Based on Access Patterns. In: *IoTBDS*. 2024. (pp. 185-190). | | | |
| **Link/DOI:** http://dx.doi.org/10.5220/0012621300003705 | | | |

| | | | |
|---|---|---|---|
| **Autori**: Tiberiu Boros, Radu Chivereanu, Octavian Purcaru | **Publicatie**: UNLP (LREC-COLING Workshop) 2024 | **Clasa**: B (A) (CORE) | **Punctaj**: 2 |
| **Titlu**: Fine-Tuning and Retrieval Augmented Generation for Question Answering Using Affordable Large Language Models. | | | |
| **Referință:** BOROȘ, Tiberiu, et al. Fine-Tuning and Retrieval Augmented Generation for Question Answering Using Affordable Large Language Models. In: *Proceedings of the Third Ukrainian Natural Language Processing Workshop (UNLP)@ LREC-COLING 2024*. 2024. (pp. 75-82). | | | |
| **Link/DOI:** https://aclanthology.org/2024.unlp-1.10.pdf | | | |

**A*+A = 12+2+8+8+8+8 = 46 (prag minim 24)**
**A*+A+B = 46+2+1+2+2+1+8 =62 (prag minim 40)**
**A*+A+B+C= 62+2+2+2+0.5+2+2+2=74.5 (prag minim 56)**

**Perspectiva C – Impactul rezultatelor**

| **Lucrarea citată** | **TUFIŞ, Dan; BOROŞ, Tiberiu; DUMITRESCU, Stefan Daniel. The RACAI speech translation system challenges of morphologically rich languages. In: *2013 7th Conference on Speech Technology and Human-Computer Dialogue (SpeD)*. IEEE, 2013. pp. 1-10. Link/DOI:** http://dx.doi.org/10.1109/SpeD.2013.6682657 | |
|---|---|---|
| **Citari** | **Clasa** | **Punctaj** |
| CIOBANU, Dragoş. Automatic speech recognition in the professional translation process. *Translation Spaces. A multidisciplinary, multimedia, and multilingual journal of translation*, 2016, 5.1: 124-144. DOI: 10.1075/bct.101.07cio | D | 1.00 |
| | | |
| **Lucrarea citată** | **BOROŞ, Tiberiu; BURTICA, Ruxandra. GBD-NER at PARSEME shared task 2018: Multi-word expression detection using bidirectional long-short-term memory networks and graph-based decoding. In: *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*. 2018. pp. 254-260. Link/DOI:** https://aclanthology.org/W18-4928.pdf | |
| **Citări** | | |
| ŠKVORC, Tadej; GANTAR, Polona; ROBNIK-ŠIKONJA, Marko. MICE: mining idioms with contextual embeddings. *Knowledge-Based Systems*, 2022, 235: 107606. DOI: 10.1016/j.knosys.2021.107606 | D | 1 |
| TASLIMIPOOR, Shiva; ROHANIAN, Omid. Shoma at parseme shared task on automatic identification of vmwes: Neural multiword expression tagging with high generalisation. *arXiv preprint arXiv:1809.03056*, 2018. DOI: 10.48550/arXiv.1809.03056 | D | 1 |
| NAGY, István; RÁCZ, Anita; VINCZE, Veronika. Detecting light verb constructions across languages. *Natural Language Engineering*, vol. 26, no. 3, ISSN *1351-3249*, pp. 319–348, 2020. | B (zona 2 in index SSCI) | 4 |

| | | |
|---|---|---|
| DOI:10.1017/S1351324919000330 | | |
| YIRMIBEŞOĞLU, Zeynep; GÜNGÖR, Tunga. ERMI at PARSEME shared task 2020: Embedding-rich multiword expression identification. In: *Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons (COLING Workshop)*. 2020. pp. 130-135. | B(A) (CORE) | 2 |
| KANCLERZ, Kamil; PIASECKI, Maciej. Deep neural representations for multiword expressions detection. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*. 2022. pp. 444-453. DOI: 10.18653/v1/2022.acl-srw.36 | A(A*) (CORE) | 6 |
| WASZCZUK, Jakub, EHREN, Rafael, STODDEN, Regina and KALLMEYER, Laura. A neural graph-based approach to verbal MWE identification. In: *Proceedings of the joint workshop on multiword expressions and WordNet (MWE-WN 2019)*. 2019. pp. 114-124. DOI: 10.18653/v1/W19-5113 | A(A*) (CORE) | 6 |
| BERK, Gözde; ERDEN, Berna; GÜNGÖR, Tunga. Representing overlaps in sequence labeling tasks with a novel tagging scheme: bigappy-unicrossy. In: *International Conference on Computational Linguistics and Intelligent Text Processing (CICLING)*. Cham: Springer Nature Switzerland, 2019. p. 622-635. DOI: 10.1007/978-3-031-24337-0_44 | B (CORE) | 4 |
| PIASECKI, Maciej; KANCLERZ, Kamil. Non-Contextual vs Contextual Word Embeddings in Multiword Expressions Detection. In: *International Conference on Computational Collective Intelligence (ICCCI)*. Cham: Springer International Publishing, 2022. pp. 193-206. DOI: 10.1007/978-3-031-16014-1_16 | B (CORE) | 4 |
| ZRIBI, Chiraz Ben Othmane. English–Arabic collocation extraction to enhance Arabic collocation identification. *Knowledge and Information Systems,* ISSN *0219-1377*, 2020, 62.6: 2439-2459. DOI: 10.1007/s10115-019-01428-0 | B (zona 2 in SCIE) | 4 |
| CIRILLO, Nicola and PAONE, Antonietta. The Role of Semi-productivity in Multiword Expression Identification: Why can BERT Capture novel MWEs?. *Computational and Corpus-based Phraseology*, 2022, 28: 152. DOI: 10.26615/978-954-452-080-9_019 | D | 1 |
| ERYIĞIT, GülŞen; ŞENTAŞ, Ali; MONTI, Johanna. Gamified crowdsourcing for idiom corpora construction. *Natural Language Engineering,* ISSN *1351-3249*, 2023, 29.4: 909-941. DOI: 10.1017/S1351324921000401 | B (zona 2 in index SSCI) | 4 |
| | | |

| Lucrarea Citată | **BOROŞ, Tiberiu; DUMITRESCU, Ştefan Daniel; BURTICA, Ruxandra. NLP-Cube: End-to-end raw text processing with neural networks. In: *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. 2018. pp. 171-179.**<br>**Link/DOI:** https://aclanthology.org/K18-2017.pdf | | |
|---|---|---|---|
| **Citări** | | | |
| AL-MOSLMI, Tareq; GALLOFRÉ OCAÑA, Marc; OPDAHL, Andreas and VERES, Csaba . Named entity extraction for knowledge graphs: A literature overview. *IEEE Access, ISSN 2169-3536*, 2020, vol. 8 pp. 32862-32881.<br>DOI: 10.1109/ACCESS.2020.2973928 | | A (zona 1 în index SCIE) | 8 |
| SEN, Ovishake; FUAD, Mohtasim; ISLAM, Nazrul; RABBI, Jakaria; MASUD, Mehedi; HASAN, Kamrul; AWAL, Abdul; AHMED FIME, Awal; HASAN FUAD, Tahmid; SIKDER, Delowar; Raihan and IFTEE, Akil (2022). Bangla natural language processing: A comprehensive analysis of classical, machine learning, and deep learning-based methods. *IEEE Access, ISSN 2169-3536*, vol. *10*, pp. 38999-39044.<br>DOI: 10.1109/ACCESS.2022.3165563 | | A (zona 1 în index SCIE) | 8 |
| SOMMERSCHIELD, Thea; ASSAEL, Yannis; PAVLOPOULOS, John;STEFANAK, Vanessa;SENIOR, Andrew; DYER, Chris;BODEL, John;PRAG, Jonathan; ANDROUTSOPOULOS, Ion; de FREITAS,Nando. (2023). Machine learning for ancient languages: A survey. *Computational Linguistics, ISSN 0891-2017*, vol. vol. *49*, no.3, pp. 703-747.<br>DOI: 10.1162/coli_a_00481 | | A (Zona 1 în index SSCI) | 8 |
| TONG, Meihan; XU, Bin; WANG, Shuai; CAO, Yixin; HOU, Lei; LI, Juanzi and XIE, Jun (2020). Improving event detection via open-domain event trigger knowledge. Association for Computational Linguistics. pp. 5887–5897<br>DOI: 10.18653/v1/2020.acl-main.522 | | A* (CORE) | 12 |
| KHAN, Bilal; SHAH, Zohaib Ali; USMAN, Muhammad; KHAN, Inayat and NIAZI, Badam. (2023). Exploring the landscape of automatic text summarization: a comprehensive survey. *IEEE Access, ISSN 2169-3536, vol. 11, pp. 109819-109840*<br>DOI: 10.1109/ACCESS.2023.3322188 | | A (zona 1 in SCIE) | 8 |
| KRISHNA, Amrith; SANTRA, Bishal; GUPTA, Ashim; SATULURI, Pavankumar and GOYAL, Pawan (2021). A graph-based framework for structured prediction tasks in Sanskrit. *Computational Linguistics, ISSN 0891-2017*, vol. *46*, no. 4, pp. 785-845.<br>DOI: 10.1162/coli_a_00390 | | A (Zona 1 în index SSCI) | 8 |

| | | |
|---|---|---|
| JAF, Sardar and CALDER, Calum. Deep learning for natural language parsing. *IEEE Access, ISSN 2169-3536*, 2019, vol. 7: pp. 131363-131373. DOI: 10.1109/ACCESS.2019.2939687 | A (zona 1 în index SCIE) | 8 |
| MITROFAN, Maria, & PĂIȘ, Vasile (2022, May). Improving Romanian BioNER using a biologically inspired system. In *Proceedings of the 21st Workshop on Biomedical Language Processing* (pp. 316-322) (Workshop ACL). DOI: 10.18653/v1/2022.bionlp-1.30 | A(A*) (CORE) | 6 |
| PĂIȘ, Vasile; ION, Radu and TUFIȘ, Dan (2020, May). A processing platform relating data and tools for Romanian language. In *Proceedings of the 1st International Workshop on Language Technology Platforms* (pp. 81-88). Link: https://aclanthology.org/2020.iwltp-1.13.pdf | D | 1 |
| WANG, Xi; BYLINSKII, Zoya; HERTZMANN, Aaron; PEPPERELL, Robert (2020). Toward quantifying ambiguities in artistic images. *ACM Transactions on Applied Perception (TAP)*, vol. *17*, no. 4, ISSN *1544-3558*, (pp. 1-10). DOI: 10.1145/3418054 | D | 1 |
| VÁRADI, Tamas; NYÉKI, Bence; KOEVA, Svetla; TADIĆ, Marko; ŠTEFANEC, Vanja; OGRODNICZUK, Maciej; NITON, Bartlomiej; PKEZIK, Piotr; MITITELU, Verginica Barbu and IRIMIA, Elena (2022, June). Introducing the CURLICAT corpora: seven-language domain specific annotated corpora from curated sources. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference (LREC)* (pp. 100-108). Link: https://aclanthology.org/2022.lrec-1.11.pdf | B (CORE2022) | 4 |
| LAUR, Sven; ORASMAA, Siim; SÄRG, Dage and TAMMO, Paul (2020, May). Estnltk 1.6: Remastered estonian nlp pipeline. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 7152-7160). Link: https://aclanthology.org/2020.lrec-1.884.pdf | C (CORE2020) | 2 |
| MITROFAN, Maria; MITITELU, Verginica Barbu; MITROFAN, Grigorina. Monero. (2019, August). Monero: a biomedical gold standard corpus for the Romanian language. In *Proceedings of the 18th BioNLP Workshop and Shared Task* (ACL Workshop) (pp. 71-79). DOI: 10.18653/v1/W19-5008 | A(A*) (CORE) | 6 |
| KOEVA, Svetla; OBRESHKOV, Nikola and YALAMOV, Martin (2020, May). Natural language processing pipeline to annotate Bulgarian legislative documents. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 6988-6994). Link: https://aclanthology.org/2020.lrec-1.863.pdf | C (CORE2020) | 2 |

| | | |
|---|---|---|
| MITITELU, Verginica Barbu and MITROFAN, Mitrofan (2020). The Romanian medical treebank-simonero. In *Proceedings of the 15th Edition of the International Conference on Linguistic Resources and Tools for Natural Language Processing–ConsILR* (pp. 7-16). | D | 1 |
| PĂIȘ, Vasile (2020). Multiple annotation pipelines inside the RELATE platform. In *Proceedings of the 15th International Conference on Linguistic Resources and Tools for Natural Language Processing* (pp. 65-75). | D | 1 |
| PĂIȘ, Vasile; ION, Radu, AVRAM, A. M., Mitrofan, Maria and TUFIȘ, Dan (2021). In-depth evaluation of Romanian natural language processing pipelines. *Romanian Journal of Information Science and Technology (ROMJIST)*, vol. *24*, no. 4 (pp. 384-401).<br>Link: https://www.romjist.ro/full-texts/paper700.pdf | D | 1 |
| ALVES, Diego; TADIĆ, Marko and BEKAVAC, Božo.. (2022, June). Multilingual comparative analysis of deep-learning dependency parsing results using parallel corpora. In *Proceedings of the BUCC Workshop within LREC 2022* (pp. 33-42).<br>Link: https://aclanthology.org/2022.bucc-1.5.pdf | C(B) (CORE2023) | 1 |
| CAN, Burcu; ALEÇAKIR, Hüseyin; MANANDHAR, Suresh and BOZŞAHIN, Cem (2022). Joint learning of morphology and syntax with cross-level contextual information flow. *Natural Language Engineering,* ISSN *1351-3249, vol. 28*, no. 6, (pp. 763-795).<br>DOI: 10.1017/S1351324921000371 | B (zona 2 in index SSCI) | 4 |
| MARTINEZ, Amaia Aguirregoitia; KORTAZAR, Kepa Bengoetxea and GONZALES-DIOS, Itziar (2021). Are CLIL texts too complicated? *Journal of immersion and content-based language education, **ISSN**: 2212-8433*, vol. *9*, no. 1, (pp. 4-30). | D | 1 |
| YOO, Inseon; KIM, Hyuntai. Created era estimation of old Korean documents via deep neural network. *Heritage Science*, 2022, vol. 10 no.1: 144<br>DOI: 10.1186/s40494-022-00772-9 | C (zona 2 in index SCIE) | 2 |
| ŠUPPA, Marek and JARIABKA, Ondrej. Benchmarking pre-trained language models for multilingual ner: Traspas at the bsnlp2021 shared task. In: *Proceedings of the 8th Workshop on Balto-Slavic Natural Language Processing (EACL Workshop)*. 2021. (pp. 105-114).<br>Link: https://aclanthology.org/2021.bsnlp-1.13.pdf | B (A) (CORE) | 2 |
| PORTA-ZAMORANO, Jordi; ESPINOSA-ANKE, Luis. Overview of CAPITEL shared tasks at IberLEF 2020: Named entity recognition and universal dependencies parsing. *arXiv preprint arXiv:2011.05932*, 2020. | D | 1 |
| BRICIU, Anamaria; CĂLIN, Alina-Delia; MIHOLCA, Diana-Lucia; MOROZ-DUBENCO, Cristiana; PETRAȘCU, Vladiela and DASCĂLU, George. (2024). Machine-Learning-Based Approaches for | D | 1 |

| | | |
|---|---|---|
| Multi-Level Sentiment Analysis of Romanian Reviews. *Mathematics, ISSN **2227-7390***, vol. *12*, no.3, (pp. 456-492).<br>DOI: 10.3390/math12030456 | | |
| LUO, Tianyi; LI, Xingyu; WANG, Hainan and LIU, Yang. (2020, January). Research replication prediction using weakly supervised learning. *In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings (EMNLP)*. (pp. 1464–1474)<br>DOI: 10.18653/v1/2020.findings-emnlp.132 | A (CORE2020) | 8 |
| Nuțu, Maria (2021). Deep learning approach for automatic Romanian lemmatization. *Procedia Computer Science, ISSN 1877-0509*, *192*, 49-58.<br>Link: https://www.sciencedirect.com/science/article/pii/S1877050921014939 | D | 1 |
| MITITELU, Verginica Barbu; IRIMIA, Elena; PĂIȘ, Vasile; AVRAM, Andrei-Marius and MITROFAN, Maria. (2022, June). Use case: Romanian language resources in the LOD paradigm. In *Proceedings of the 8th Workshop on Linked Data in Linguistics within the 13th Language Resources and Evaluation Conference* (pp. 35-44).<br>Link: https://aclanthology.org/2022.ldl-1.5.pdf | C (B) (CORE 2023) | 1 |
| PĂIȘ, Vasile; BARBU MITITELU, Verginica; IRIMIA, Elena; ION, Radu and TUFIȘ, Dan. (2024). Under-Represented Speech Dataset from Open Data: Case Study on the Romanian Language. *Applied Sciences, ISSN 2076-3417*, vol. *14*, no. 19, 9043.<br>DOI: 10.3390/app14199043 | D | 1 |
| Hládek, Daniel; Harahus, Maros; Staš, Jan and Pleva, Matus (2023). Slovak Language Models for Basic Preprocessing Tasks in Python. *Journal of Linguistics/Jazykovedný casopis, **ISSN:** 1338-4287*, vol. *74*, no. 1, (pp. 323-332)<br>DOI: 10.2478/jazcas-2023-0049 | D | 1 |
| TANGUY, Ludovic; BRUNET, Pauline and FERRET, Olivier. (2020, May). Extrinsic Evaluation of French Dependency Parsers on a Specialized Corpus: Comparison of Distributional Thesauri. In the *12th International Conference on Language Resources and Evaluation (LREC 2020)* (pp. 5822-5830). | C (CORE) | 2 |
| Papadopoulou, A. (2019). Automatic Error Detection and Correction in Neural Machine Translation: A comparative study of Swedish to English and Greek to English. | D | 1 |
| LEÓN, Fernando Sánchez. Combining Different Parsers and Datasets for CAPITEL UD Parsing. In *IberLEF@ SEPLN*. 2020. p. 39-44. | D | 1 |
| MANOLACHE, Cristian; CUCU, Horia; BURILEANU, Corneliu. Lemma-based dynamic time warping search for keyword spotting applications in Romanian. In: *2019 International Conference on Speech* | D | 1 |

| | | |
|---|---|---|
| *Technology and Human-Computer Dialogue (SpeD)*. IEEE, 2019. p. 1-9. DOI: 10.1109/SPED.2019.8906586 | | |
| Wawer, Aleksander (2021, September). ComboNER: A Lightweight All-In-One POS Tagger, Dependency Parser and NER. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)* (pp. 1508-1514). DOI: 10.26615/978-954-452-072-4_169 | (National) (CORE 2021) | 2 |
| PADURARU, Ciprian; PADURARU, Miruna and PATILEA, Catalina Camelia (2021, November). Task distribution and human resource management using reinforcement learning. In *2021 36th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW)* (pp. 96-101). IEEE. DOI: 10.1109/ASEW52652.2021.00029 | D | 1 |
| PĂIȘ, Vasile; MITITELU, Verginica Barbu; IRIMIA, Elena, ION, Radu; BADEA, Valentin; TUFIȘ, Dan, and STEREA, Grosu Cosmin. Towards Anonymising The Romanian Jurisprudence. *Linguistic Resources And Tools For Natural Language Processing*, (pp. 85-99) | D | 1 |
| | | |
| **Lucrarea citată** | **BOROȘ, Tiberiu; PIPA, Sonia; MITITELU, Verginica Barbu and TUFIȘ, Dan (2017, April). A data-driven approach to verbal multiword expression detection. PARSEME Shared Task system description paper. In *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)* (pp. 121-126). Link/DOI: https://aclanthology.org/W17-1716.pdf** | |
| **Citări** | | |
| NAGY, István; RÁCZ, Anita and VINCZE, Veronika (2020). Detecting light verb constructions across languages. *Natural Language Engineering,* ISSN *1351-3249*, *26*(3), 319-348. DOI: 10.1017/S1351324919000330 | B (zona 2 in index SSCI) | 2 |
| INURRIETA, Urmeneta Usoa; ADURIZ, Itziar; DIAZ DE ILARRAZA SANCHEZ, Arantza; LABAKA INTXAUSPE, Gorka and SARASOLA GABIOLA, Kepa Mirena. (2020). Learning about phraseology from corpora: A linguistically motivated approach for Multiword Expression identification. *Plos one, ISSN 1932-6203*, vol. *15*, no. 8. DOI: 10.1371/journal.pone.0237767 | D | 0.5 |
| AL SAIED, Hazem; CANDITO, Marie and CONSTANT, Matthieu (2018). A transitionbased verbal multiword expression analyzer. In *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop* (Vol. 2, p. 209). Language Science Press. | D | 0.5 |

| | | |
|---|---|---|
| RAMISCH, Carlos; WALSH, Abigail; BLANCHARD, Thomas and TASLIMIPOOR, Shiva (2023, May). A survey of MWE identification experiments: the devil is in the details. In *Proceedings of the 19th workshop on multiword expressions (MWE 2023) (EACL Workshop)* (pp. 106-120). DOI: 10.18653/v1/2023.mwe-1.15 | B(A) (CORE) | 1 |
| EHREN, Rafael; LICHTE, Timm and SAMIH, Younes (2018, August). Mumpitz at PARSEME shared task 2018: A bidirectional LSTM for the identification of verbal multiword expressions. In *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018) (EACL Workshop)* (pp. 261-267). Link: https://aclanthology.org/W18-4929.pdf | B(A) (CORE) | 1 |
| MALDONADO, Alfredo and QASEMIZADEH, Behrang.(2018). Analysis and Insights from the PARSEME Shared Task dataset. In *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop* (Vol. 2, p. 149). Language Science Press. | D | 0.5 |
| SIMKÓ, Katalin Ilona; KOVÁCS, Viktória and VINCZE, Veronika.(2018). Identifying verbal multiword expressions with POS tagging and parsing techniques. In *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop* (Vol. 2, p. 227). Language Science Press. | D | 0.5 |
| HERRERA, Antonio Jesús Tamayo (2024). Hacia la estandarización de la evaluación en la extracción de términos1. In *Terminología del español: el término/Spanish Terminology: The Term* (pp. 223-242). Routledge. | D | 0.5 |
| VARIS, Dusan and KLYUEVA, Natalia. (2018, May). Improving a neural-based tagger for multiword expressions identification. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Link: https://aclanthology.org/L18-1401.pdf | C | 1 |
| MONTI, Johanna; CORDEIRO, Silvio; RAMISCH, Carlos; SANGATI, Federico; SAVARY, Agata; VINCZE, Veronika. (2018, December). Advances in Multiword Expression Identification for the Italian language: The PARSEME shared task edition 1.1. In *Fifth Italian Conference on Computational Linguistics (CLiC-it 2018)*. | D | 0.5 |
| | | |
| **Lucrarea citată** | **BOROȘ, Tiberiu and TUFIȘ, Dan. Romanian-English Speech Translation. Proceedings of the Romanian Academy, Series A, Volume 15, Number 1/2014, ISSN, 1454-9069. (pp. 68–75) Link/DOI: https://academiaromana.ro/sectii2002/proceedings/doc2014-1/09-Tufis.pdf** | |
| **Citări** | | |
| TODEREAN, Gavril; BUZA, Ovidiu and DOMOKOS, József. (2015, October). Achievements in the field of voice synthesis for Romanian. In Speech Technology and Human-Computer Dialogue (SpeD), 2015. (pp. | D | 1.00 |

| | | |
|---|---|---|
| 1-7). IEEE.<br>DOI: 10.1109/SPED.2015.7343078 | | |
| | | |
| **Lucrarea citată** | **BOROȘ, Tiberiu; STAN, Adriana, WATTS, Oliver and DUMITRESCU, Stefan (2014). RSS-TOBI-A Prosodically Enhanced Romanian Speech Corpus. Proceedings of 9th LREC 2014, (pp. 316-320).<br>Link/DOI: https://www.cstr.ed.ac.uk/downloads/publications/2014/boros-2014.pdf** | |
| **Citări** | | |
| DALL, Rasmus; GONZALVO, Xavi (2016). JNDSLAM: A SLAM extension for speech synthesis. In Proc. Speech Prosody, ISSN 2333-2042 (pp. 1024-1028).<br>DOI: 10.21437/SpeechProsody.2016-210 | D | 0.50 |
| SUCIU, George; TOMA, Ștefan-Adrian and CHEVERESAN, Romulus. Towards a Continuous Speech Corpus for Banking Domain Automatic Speech Recognition. In Speech Technology and Human-Computer Dialogue (SpeD)<br>DOI: 10.1109/SPED.2017.7990436 | D | 0.50 |
| | | |
| **Lucrarea citată** | **DUMITRESCU, Stefan; BOROȘ, Tiberiu and ION, Radu (2014). Crowd-Sourced, Automatic Speech-Corpora Collection–Building the Romanian Anonymous Speech Corpus. CCURL 2014: Collaboration and Computing for Under-Resourced Languages in the Linked Open Data Era, (pp. 90-94).<br>Link/DOI: http://www.syros.aegean.gr/users/spyrosv/papers/ccurl14.pdf#page=97** | |
| **Citări** | | |
| STAN, Adriana; DINESCU, Florina; ȚIPLE, Cristina; MEZA, Șerban; ORZA, Bogdan; CHIRILĂ, Magdalena and GIURGIU, Mircea. (2017, July). The SWARA speech corpus: A large parallel Romanian read speech dataset. In *2017 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)* (pp. 1-6). IEEE.<br>DOI: 10.1109/SPED.2017.7990428 | D | 1 |
| Avram, A. M., Nichita, M. V., Bartusica, R. G., & Mihai, M. V. (2022, June). RoSAC: A speech corpus for transcribing Romanian emergency calls. In *2022 14th International Conference on Communications (COMM)* (pp. 1-5). IEEE.<br>DOI: 10.1109/COMM54429.2022.9817214 | D | 1 |
| | | |

| | | |
|---|---|---|
| **Lucrarea citată** | **BOROȘ, Tiberiu; DUMITRESCU, Ștefan; ZAFIU, Adrian; MITITELU, Verginica and VĂDUVA, Ionut (2014, June). RACAI GEC–a hybrid approach to grammatical error correction. In *Proceedings of the Eighteenth Conference on Computational Natural Language Learning (CONLL): Shared Task* (pp. 43-48). Link/DOI: 10.3115/v1/W14-1705** | |
| **Citări** | | |
| NG, Hwee Tou; WU, Siew Mei; BRISCOE, Ted; HADIWINOTO, Christian; SUSANTO, Raymond Hendy and BRYANT, Christopher. (2014, May). The CoNLL-2014 Shared Task on Grammatical Error Correction. In CoNLL Shared Task (pp. 1-14). DOI: 10.3115/v1/W14-1701 | A (CORE) | 2.00 |
| FELICE, Mariano. (2016). Artificial error generation for translation-based grammatical error correction (No. UCAM-CL-TR-895). University of Cambridge, Computer Laboratory. Link: https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-895.pdf | D | 0.25 |
| DRONEN, Nicholas. (2016). Correcting Writing Errors with Convolutional Neural Networks (Doctoral dissertation, University of Colorado at Boulder). | D | 0.25 |
| | | |
| **Lucrarea citată** | **BOROȘ, Tiberiu and DUMITRESCU, Ștefan. (2015, October). Robust deep-learning models for text-to-speech synthesis support on embedded devices. In Proceedings of the 7th International Conference on Management of computational and collective intElligence in Digital EcoSystems (pp. 98-102). ACM. Link/DOI: 10.1145/2857218.2857234** | |
| **Citări** | | |
| YANG, Chaofei; WU, Qing; LI, Hai and CHEN, Yiran. (2017). Generative Poisoning Attack Method Against Neural Networks. arXiv preprint arXiv:1703.01340. Link: https://arxiv.org/pdf/1703.01340 | D | 1.00 |
| SUCIU, George; DOBRE, Robert Alexandru; BUTCA, Cristina; SUCIU, Victor; MIHAILA, Ioana and CHEVERESAN, Romulus. (2016, June). Search based applications for speech processing. In Electronics, Computers and Artificial Intelligence (ECAI), 2016 8th International Conference on (pp. 1-6). IEEE. DOI: 10.1109/ECAI.2016.7861101 | D | 1.00 |
| | | |
| **Lucrarea citată** | **BOROȘ, Tiberiu; ION, Radu and TUFIȘ, Dan (2013, August). Large tagset labeling using Feed Forward Neural Networks. Case study on Romanian Language. In ACL (1) (pp.** | |

| | | |
|---|---|---|
| | **692-700).**<br>**Link/DOI: https://aclanthology.org/P13-1068.pdf** | |
| **Citări** | | |
| EGER, Steffen; VOR DER BRÜCK, Tim; MEHLER, Alexander.. (2015). Lexicon-assisted tagging and lemmatization in Latin: A comparison of six taggers and two lemmatization methods. LaTeCH 2015, 105. (Worskhop ACL-IJCNLP). (pp. 105-113)<br>Link: https://aclanthology.org/W15-3716.pdf | A(A*) (CORE) | 6.00 |
| Dumitrescu, Stefam. (2017, July). Cassandra smart-home system description. In *2017 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)* (pp. 1-6). IEEE.<br>DOI: 10.1109/SPED.2017.7990440 | D | 1.00 |
| | | |
| **Lucrarea citată** | **BOROȘ, Tiberiu. (2013). A unified lexical processing framework based on the Margin Infused Relaxed Algorithm. A case study on the Romanian Language. In RANLP (pp. 91-97).**<br>**Link/DOI: https://aclanthology.org/R13-1012.pdf** | |
| **Citări** | | |
| TOMA, Ștefan-Adrian; STAN, Adriana; PURA, Mihai-Lică and BÂRSAN, Traian. MaRePhoR–An Open Access Machine-Readable Phonetic Dictionary for Romanian. 2017 International Conference on Speech Technology and Human-Computer Dialogue (SpeD). IEEE (pp. 1-6)<br>DOI: 10.1109/SPED.2017.7990435 | D | 1.00 |
| DUMITRESCU, Ștefan. Cassandra Smart-home System Description. 2017 International Conference on Speech Technology and Human-Computer Dialogue (SpeD). IEEE<br>DOI: 10.1109/SPED.2017.7990440 | D | 1.00 |
| | | |
| **Lucrarea citată** | **BOROȘ, Tiberiu; ȘTEFĂNESCU, Dan and ION, Radu (2012). Bermuda, a data-driven tool for phonetic transcription of words. In Natural Language Processing for Improving Textual Accessibility (NLP4ITA) (LREC Workshop) (p. 35).**<br>**Link/DOI:**<br>**https://www.academia.edu/download/98806157/25.NLP4ITA-Proceedings.pdf#page=40** | |
| **Citări** | | |

| | | |
|---|---|---|
| DOMOKOS, József; BUZA, Ovidiu and TODEREAN, Gavril. Romanian phonetic transcription dictionary for speeding up language technology development. *Language Resources and Evaluation*, 2015, vol. 49, (pp. 311-325).<br>DOI: 10.1007/s10579-013-9262-z | C | 2 |
| DIACONESCU, Ștefan-Stelian; RIZEA, Monica-Mihaela; CODÎRLAȘU, Felicia-Carmen; IONESCU, Mihaela; RĂDULESCU, Monica; MINCĂ, Andrei and FULEA, Ștefan. (2015, October). Methods for automatic generation of GRAALAN-based phonetic databases. In Speech Technology and Human-Computer Dialogue (SpeD), 2015 International Conference on (pp. 1-8). IEEE.<br>DOI: 10.1109/SPED.2015.7343082 | D | 1.00 |
| | | |
| **Lucrarea citată** | **TUFIȘ, Dan; BOROȘ, Tiberiu and DUMITRESCU, Ștefan. (2013, October). The RACAI Speech Translation System. In Proceedings of the 7th International Conference on Speech Technology and Human-Computer Dialogue (SPED 2013). Cluj-Napoca.** | |
| **Citări** | | |
| TODEREAN, Gavril; BUZA, Ovidiu and DOMOKOS, József. (2015, October). Achievements in the field of voice synthesis for Romanian. In Speech Technology and Human-Computer Dialogue (SpeD), 2015 International Conference on (pp. 1-7). IEEE.<br>DOI: 10.1109/SPED.2015.7343078 | D | 1.00 |
| IRIMIA, Elena (2015). Accelerating the Development of a Digital Corpus Annotated with Dependencies Using Resources and Tools Dedicated to Other Languages. Revista Română de Informatică și Automatică, ISSN 1220-1758 , vol. 25, nr. 3, 2015 (pp. 5-16)<br>Link: https://rria.ici.ro/documents/289/03-art-1-IRIMIA_ELENA1-OK-1.pdf | D | 1.00 |
| | | |
| **Lucrarea citată** | **BOROȘ, Tiberiu; Ștefănescu, D., & Ion, R. (2013). Handling Two Difficult Challenges for Text-to-Speech Synthesis Systems: Out-of-Vocabulary Words and Prosody: A Case Study in Romanian. In Where Humans Meet Machines (pp. 137-161). Springer New York.** | |
| **Citări** | | |
| DUMITRESCU, Ștefan. (2015). Rowordnetlib–the first API for the Romanian wordnet. Proceedings of the Romanian Academy Series A-Mathematics Physics Technical Sciences Information Science, ISSN *1454-9069* , 16(1), 87-94.<br>Link: https://acad.ro/sectii2002/proceedings/doc2015-1/12-Dumitrescu.pdf | C (Zona 3 in index SCIE) | 2.00 |

| | | |
|---|---|---|
| **Lucrarea citată** | **ION, Radu; TUFIȘ, Dan; BOROȘ, Tiberiu; CEAUȘU, Alexandru and ȘTEFĂNESCU, Dan (2010, October). On-line compilation of comparable corpora and their evaluation. In Proceedings of the 7th International Conference of Formal Approaches to South Slavic and Balkan Languages (FASSBL7) (pp. 29-34). Link/DOI: http://dcl.bas.bg/wp-content/uploads/2015/08/FASSBL7_2010_proceedings-2.pdf#page=30** | |
| **Citări** | | |
| MASTROPAVLOS, Nikos; PAPAVASSILIOU, Vassilis. (2011). Automatic acquisition of bilingual language resources. In Proceedings of the 10th International Conference of Greek Linguistics, Komotini, Greece. Link: http://nlp.ilsp.gr/nlp/ICGL2011_Mastropavlos_Papavassiliou.pdf | D | 0.33 |
| STAMBOLIEVA, Ekaterina. (2013, August). Learning Comparable Corpora from Latent Semantic Analysis Simplified Document Space. In Proceedings of the Sixth Workshop on Building and Using Comparable Corpora (BUCC 2013) at ACL (pp. 129-137). Link: https://aclanthology.org/W13-2516.pdf | A(A*) (CORE) | 2 |
| SAAD, Motaz. (2015). Mining Documents and Sentiments in Cross-lingual Context (Doctoral dissertation, Université de Lorraine). Link: https://inria.hal.science/tel-01751251/file/saad_phd.pdf | D | 0.33 |
| STAMBOLIEVA, Ekaterina (2013). Reporting Preliminary Automatic Comparable Corpora Compilation Results. In Recent Advances in Natural Language Processing (RANLP) (pp. 149-156). Link: https://aclanthology.org/R13-2022.pdf | C (CORE) | 0.67 |
| YAPOMO, Manuela (2013). Construction de corpus multilingues: état de l'art. TALN-RÉCITAL 2013, 56. Link: https://hal.science/hal-01073648/file/Article%2Bposter_Recital.pdf | C | 0.67 |

**A*+A+B = 138 (prag minim 40)**
**A*+A+B+C+D =138+37=175 (prag minim 120) (din lista anterioară - nu include toate lucrările) - estimat 250 (fără autocitări directe/indirecte)**

**Perspectiva D – Performanță academică**

**i. Cărți autor/editate și capitole publicate în edituri de categoria (conform clasamentului SENSE): total 8.67 puncte**

| Titlu | Titlu carte | Autori | Editură | ISBN | Punctaj |
|---|---|---|---|---|---|
| Improving the RACAI Neural Network MSD Tagger | Engineering Applications of Neural Networks | Tiberiu Boroș, Ștefan Daniel Dumitrescu | Springer | 978-3-642-4101 3-0 | 2 |
| Filtering, normalizing and spell-checking the Romanian N-grams from Google's Web 1TB 5-Grams | Towards Multilingual Europe 2020: A Romanian perspective | Dan Tufiș, Radu Ion, Tiberiu Boroș | Editura Academiei Române | 978-973-27-228 2-4 | 2 |
| Experiments on Language and Translation Models Adaptation for Statistical Machine Translation | Towards Multilingual Europe 2020: A Romanian perspective | Ștefan Dumitrescu, Radu Ion, Dan Ștefănescu, Tiberiu Boroș, Dan Tufiș | Editura Academiei Române | 978-973-27-228 2-4 | 0.67 |
| RACAI NLP Tools for Text-to-Speech Synthesis | Towards Multilingual Europe 2020: A Romanian perspective | Tiberiu Boroș | Editura Academiei Române | 978-973-27-228 2-4 | 2 |
| Handling Two Difficult Challenges for Text-to-Speech Synthesis Systems: Out-of-Vocabulary Words and Prosody: A Case Study in Romanian | Where Humans Meet Machines Innovative Solutions for Knotty Natural-Language Problems | Tiberiu Boroș, Dan Ștefănescu, Radu Ion | Springer | 978-1-4614-693 3-9 | 2 |

**v. Director (coordonator/responsabil) | membru al unui grant/proiect/contract/program de cercetare național/internațional: total 24 puncte**

| Denumire proiect | Rol | Punctaj |
|---|---|---|
| Accurat | Executant (membru) | 4 |
| Metanet | Executant (membru) | 4 |
| ANVSIB | Executant (membru) | 4 |
| SIAFIM | Executant (membru) | 4 |
| HeimdallR (PED 229/2018) | **Director** (https://heimdall.racai.ro/)<br>(https://uefiscdi.gov.ro/resource-8201) (p. 7)<br>(https://uefiscdi.gov.ro/resource-87562) (p. 10) | 8 |

**vi. Membru în comitetul științific (8 puncte)**

a. 18th Workshop on Multiword Expressions (MWE 2022)[1] - (LREC workshop) C - 2 puncte
b. 19th Workshop on Multiword Expressions (MWE 2023)[2] - (EACL Workshop) - B - 4 puncte
c. 20th Workshop on Multiword Expressions (MWE 2024)[3] - (LREC-Coling workshop) - C - 2 puncte

**viii. Keynote/invited speaker la evenimente/universități: total 5 puncte**

a. EUROLAN[4] (școală de vară internațională): 4 puncte
b. SPED 2013[5] (invited paper): 1 punct

**xii. Membru în comisii de îndrumare:** 1 punct

a. Membru în comisia de îndrumare pentru Constantin Nicolae.

**xiii. Brevete şi invenţii active (OSIM, ORDA etc.): total 42 puncte**

a. SIAFIM MAPP – INT. nr. RG II/6137/29.07.2015 – IES. nr. RG II/6137/21.08.2015 (12 puncte)
b. SIAFIM EI – INT. nr. RG II/6138/29.07.2015 – IES. nr. RG II/6138/21.08.2015 (12 puncte)
c. US Patent App. 18/163,170 - Voice audio compression using neural networks (12/4 - 3 puncte)

---

[1] https://multiword.org/mwe2022/#committee
[2] https://multiword.org/mwe2023/#committee
[3] https://multiword.org/mweud2024/#committee
[4] http://eurolan.info.uaic.ro/2015/lecturers/
[5] http://www.sped2013.ro/index.php?page=scientific-program - The RACAI Speech Translation System. Challenges of Morphologically Rich Languages

d. US Patent 11,816,210 - Risk-based alerting for computer security (12/4 - 3 puncte)
e. US Patent 11,146,580 - Script and command line exploitation detection (12 puncte)

**xiii. Dezvoltarea de pachete şi instrumente software, dezvoltarea de resurse şi colecţii de date de largă utilitate (probate prin numărd de accccesări, publicarea pe site-uri open source etc.): 2*(11-2) = 18 puncte**

| Denumire | Autori | Descriere | Adresă | Punctaj |
|---|---|---|---|---|
| RASC | Tiberiu Boroș, Ștefan Dumitrescu | Platformă achiziție date audio prin crowd-sourcing | http://rasc.racai.ro (scos din folosință) | |
| MLPLA | Tiberiu Boroș, Ștefan Dumitrescu | Sistem modular pentru prelucrare text | http://slp.racai.ro/index.php/mlpla/ http://slp.racai.ro/index.php/mlpla-new/ | |
| SSLA | Tiberiu Boroș, Ștefan Dumitrescu | Sintetizator de voce cu suport pentru sinteză parametrică statistică și sinteză concatenativă | http://slp.racai.ro/index.php/ssla/ | |
| RACAI Translation System | Tiberiu Boroș, Ștefan Dumitrescu | Sistem de traducere automată cu suport pentru conversie text în sunet | http://www.racai.ro/tools/translation/racai-translation-system/ | |
| NLP-Cube | Tiberiu Boros, Ștefan Dumitrescu | Unelte pentru procesarea textului - suport pentru 60 de limbi (tokenizare, part-of-speech tagging, | https://github.com/adobe/NLP-Cube/ | |
| TTS-Cube | Tiberiu Boros | End-to-end speech synthesis tools based on generative adversarial training | https://github.com/tiberiu44/TTS-Cube | |
| OSAS | Tiberiu Boros, Andrei Cotaie | One stop anomaly shop - statistical modeling of log data | https://github.com/adobe/OSAS | |
| Stringlifier | Tiberiu Boros, Andrei Cotaie | Character level string labeling for JWT and random tokens | https://github.com/adobe/stringlifier | |
| Tripod | Tiberiu Boros | Tool for computing latent representations for text | https://github.com/adobe/tripod | |

| | | | | |
|---|---|---|---|---|
| libLOL | Tiberiu Boros, Andrei Cotaie | Machine Learning tool for detecting Living off the Land Attacks | https://github.com/adobe/liblol | |
| Sherlock/UNLP | Tiberiu Boros, Radu Chivereanu | Question Answering system using in the UNLP Shared Task | https://github.com/adobe/sherlock-backend/tree/UNLP2024 | |

**xvii. Premii şi alte merite (la decizia universităţii sau institutului de cercetare): total 3.6 puncte** – cf. deciziei conducerii ICIA din 26 iunie 2017

a. Premiul „Mihai Drăgănescu" – Academia Română
b. Microsoft Research Award

**Punctaj total: 110,27 (prag minim 60)**
**Condiția satisfăcută pentru un grant: PED 229/2018**