



ACADEMIA ROMÂNĂ
SCOSAAR

Anexa nr.3

AVIZAT,

Director ȘCOALA DOCTORALĂ.....

1. Îndeplinirea standardelor IOSUD superioare standardelor minimale naționale* DA | NU

2. Îndeplinirea standardelor IOSUD egale standardelor minimale naționale* DA | NU

FIȘA DE ÎNDEPLINIRE A STANDARDDELOR IOSUD

FIȘA DE VERIFICARE a îndeplinirii standardelor IOSUD

Candidat: PĂIS, VASILE-FLORIAN

Data: 14.01.2025

Semnătura:

*se va alege una dintre variante

**Autoevaluarea activității profesionale de cercetare și a recunoașterii impactului acesteia, conform
Anexei nr. 15 la OMEC 6129/2016**

Etica cercetării

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

Activitatea didactică și profesională (A1)

A1.1.1 Cărți de autor sau capitole de specialitate la edituri cu ISBN Internaționale

Titlu	Autori/Editor	Editură	ISBN	Categoria SENSE	Bibl. World Cat	Punctaj
Speech Recognition Technology and Applications https://novapublishers.com/shop/speech-recognition-technology-and-applications/	Carte de specialitate Păiș, Vasile	Nova Science Publishers	978-1-68507-929-1	C	31	50
Self-supervised pre-training in speech recognition systems	Avram, Andrei-Marius and Păiș, Vasile and Tufiș, Dan	Nova Science Publishers	978-1-68507-929-1	C		4.17
The impact of speech recognition performance on human-computer interaction	Păiș, Vasile and Avram, Andrei-Marius and Ion, Radu and Tufiș, Dan	Nova Science Publishers	978-1-68507-929-1	C		3.13
Punctuation recovery for Romanian transcribed documents	Păiș, Vasile	Nova Science Publishers	978-1-68507-929-1	C		12.5
Linguistic linked open data for speech processing	Păiș, Vasile and Barbu-Mititelu, Verginica	Nova Science Publishers	978-1-68507-929-1	C		6.25
Language Report Romanian. Chapter in Georg Rehm and Andy Way, editors. European Language Equality: A Strategic	Păiș, Vasile and Tufiș, Dan	Springer	978-3-031-28818-0	B	256	12.5

Agenda for Digital Language Equality. Cognitive Technologies. Springer, June 2022.							
MARCELL – A project to remember: hard work of a friendly consortium under wise coordination http://real.mtak.hu/126576/1/VT70_kotet_honlapra.pdf#page=130	Tufiş, Dan and Păiș, Vasile and Barbu Mititelu, Verginica and Ion, Radu and Irimia, Elena and Avram, Andrei and Curea, Eric	Nyelvtudományi Kutatóközpont, Budapest	978-963-9074-90-3				1.79
RELATE: a modern processing platform for Romanian language. Chapter in Tools, Methods, and Solutions for the Exploration of Romanian Corpora. Institutul European. https://www.euroinst.ro/titlu.php?id=1813	Păiș, Vasile and Ion, Radu and Avram, Andrei and Mitrofan, Maria and Tufiş, Dan	Institutul European	978-606-24-0393-9				2
Longitudinal learner corpus of Romanian (Long_LECOR)	Carte de specialitate Mîrzea Vasile, Carmen, Elena Irimia, Vasile Păiș	Editura Universității din București	978-606-16-1555-1 (in curs de aparitie)				13.34

Total A1.1.1 : 105.68

Total A1. 105.68

A2 Activitatea de cercetare

A2.1 Articole în reviste cotate ISI, și lucrări în volumele unor manifestări științifice indexate ISI

Autori: Vasile Păiș , Verginica Barbu Mititelu, Elena Irimia, Radu Ion, Dan Tufiş	Publicație: Applied Sciences 2024	IF: 2.5	Punctaj: 20
Titlu: Underrepresented Speech Dataset from Open Data: Case Study on the Romanian Language			
Link: https://www.mdpi.com/2076-3417/14/19/9043			

Autori: <u>Vasile Păis</u> , Radu Ion, Elena Irimia, Verginica Barbu Mititelu, Valentin Badea, Dan Tufiş	Publicație: Artificial Intelligence and Law 2024	IF: 3.1	Punctaj:19.67
Titlu: System for the anonymization of Romanian jurisprudence Link: https://link.springer.com/article/10.1007/s10506-024-09420-y			
Autori: <u>Vasile Păis</u> , Maria Mitrofan, Carol Luca Gasan, Alexandru Ianov, Corvin Ghiță, Vlad Silviu Coneschi, Andrei Onuț	Publicație: SemanticWeb Journal 2024	IF: 3.1	Punctaj:12.43
Titlu: LegalNERo: A linked corpus for named entity recognition in the Romanian legal domain Link: https://content.iopress.com/articles/semantic-web/sw233351			
Autori: Avram, Andrei-Marius and Mititelu, Verginica Barbu and <u>Păis, Vasile</u> and Cercel, Dumitru-Clementin and Trăuşan-Matu, Ştefan	Publicație: Mathematics 2023	IF: 2.3	Punctaj:14.2
Titlu: Multilingual Multiword Expression Identification Using Lateral Inhibition and Domain Adaptation Link: https://www.mdpi.com/2227-7390/11/11/2548			
Autori: <u>Păis, Vasile</u> and Ion, Radu and Avram, Andrei-Marius and Mitrofan, Maria and Tufiş, Dan	Publicație: ROMJIST 2021	IF: 3.7	Punctaj:19.8
Titlu: In-depth evaluation of Romanian natural language processing pipelines Link: https://www.romjist.ro/full-texts/paper700.pdf			
Autori: <u>Păis, Vasile</u> and Tufiş, Dan	Publicație: Artificial Intelligence Review 2022	IF: 10.7	Punctaj: 119.5
Titlu: Capitalization and punctuation restoration: a survey Link: https://link.springer.com/article/10.1007/s10462-021-10051-x			
Autori: Avram, Andrei-Marius and <u>Păis, Vasile</u> and Tufiş, Dan	Publicație: Proceedings of the Romanian Academy Series A, 2020	IF: 0.7	Punctaj: 13.0

Titlu: Towards a Romanian end-to-end automatic speech recognition based on DeepSpeech2
Link: https://acad.ro/sectii2002/proceedings/doc2020-4/11-Avram_Tufis.pdf

Autori: <u>Păis, Vasile</u> and Tufiş, Dan	Publicație: Proceedings of the Romanian Academy Series A, 2018	IF: 0.7	Punctaj: 19.5
Titlu: Computing distributed representations of words using the CoRoLa corpus Link: https://academiaromana.ro/sectii2002/proceedings/doc2018-2/Art12Pais.pdf			

Autori: Andrei-Marius Avram, Andreea Iuga, George-Vlad Manolache, Vlad-Cristian Matei, Răzvan-Gabriel Micliuş, Vlad-Andrei Muntean, Manuel-Petru Sorlescu, Drago-Andrei Şerban, Adrian-Dinu Urse, Vasile Păiş , Dumitru-Clementin Cercel	Publicație: ICDAR 2024	IF: 0.25	Punctaj: 2.96
Titlu: HistNERo: Historical Named Entity Recognition for the Romanian Language Link: https://link.springer.com/chapter/10.1007/978-3-031-70543-4_8			

Autori: Carol Luca Gasan, Vasile Păiş	Publicație: IWSLT ACL 2024	IF: 0.25	Punctaj: 16.25
Titlu: Multi-Model System for Effective Subtitling Compression Link: https://aclanthology.org/2024.iwslt-1.9.pdf			

Autori: Sara Niță, Vasile Păiş	Publicație: SemEval NAACL 2024	IF: 0.25	Punctaj: 16.25
Titlu: RACAI at SemEval-2024 Task 10: Combining algorithms for code-mixed Emotion Recognition in Conversation Link: https://aclanthology.org/2024.semeval-1.149.pdf			

Autori: Vasile Păis, Dan Tufiş, Elena Irimia, Verginica Barbu Mititelu	Publicație: LAW EACL 2024	IF: 0.25	Punctaj: 8.13
Titlu: Building a corpus for the anonymization of Romanian jurisprudence			
Link: https://aclanthology.org/2024.law-1.7.pdf			

Autori: Vasile Păis	Publicație: CASE EACL 2024	IF: 0.25	Punctaj: 32.5
Titlu: RACAI at ClimateActivism 2024: Improving Detection of Hate Speech by Extending LLM Predictions with Handcrafted Features			
Link: https://aclanthology.org/2024.case-1.8.pdf			

Autori: Barbu Mititelu, Verginica and Irimia, Elena and Păis, Vasile and Avram, Andrei-Marius and Mitrofan, Maria	Publicație: LDL LREC 2022	IF: 0.25	Punctaj: 6
Titlu: Use case: Romanian language resources in the LOD paradigm			
Link: http://www.lrec-conf.org/proceedings/lrec2022/workshops/LDL/2022.ldl2022-1.0.pdf#page=43			

Autori: Păis, Vasile	Publicație: SemEval NAACL 2022	IF: 0.25	Punctaj: 30
Titlu: RACAI at SemEval-2022 Task 11: Complex named entity recognition using a lateral inhibition mechanism			
Link: https://aclanthology.org/2022.semeval-1.215			

Autori: Păis, Vasile and Mitrofan, Maria and Barbu Mititelu, Verginica and Irimia, Elena and Micu, Roxana and Gasan, Carol Luca	Publicație: CMLC LREC 2022	IF: 0.25	Punctaj: 5
Titlu: Challenges in Creating a Representative Corpus of Romanian Micro-Blogging Text			
Link: http://www.lrec-conf.org/proceedings/lrec2022/workshops/CMLC10/2022.cmlc10-1.0.pdf#page=9			

Autori: Váradi, Tamás and Nyéki, Bence and Koeva, Svetla and Tadić, Marko and Štefanec, Vanja and Ogrodniczuk, Maciej and Niño, Bartłomiej and Pezik, Piotr and Barbu Mititelu, Verginica and Irimia, Elena and Mitrofan, Maria and Păis, Vasile and Tufiş, Dan and Garabík, Radovan and Krek, Simon and Repar, Andraž	Publicație: LREC 2022	IF: 0.25	Punctaj: 1.88
---	------------------------------	-----------------	----------------------

Titlu: Introducing the CURLICAT Corpora: Seven-language Domain Specific Annotated Corpora from Curated Sources
Link: <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.11.pdf>

Autori: Avram, Andrei-Marius and Catrina, Darius and Cercei, Dumitru-Clementin and Dascalu, Mihai and Rebedea, Traian and **Păis, Vasile** and Tufiş, Dan
Titlu: Distilling the Knowledge of Romanian BERTs Using Multiple Teachers
Link: <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.39.pdf>

Autori: Mitrofan, Maria and **Păis, Vasile**
Titlu: Improving Romanian BioNER Using a Biologically Inspired System
Link: <https://aclanthology.org/2022.bionlp-1.30/>

Autori: Avram, Andrei-Marius and **Păis, Vasile** and Mitrofan, Maria
Titlu: RACAI@SMM4H'22: Tweets Disease Mention Detection Using a Neural Lateral Inhibitory Mechanism
Link: <https://aclanthology.org/2022.smm4h-1.1/>

Autori: **Păis, Vasile** and Mitrofan, Maria and Avram, Andrei-Marius
Titlu: Romanian Language Translation in the RELATE Platform
Link: <https://aclanthology.org/2022.loresmt-1.9/>

Autori: **Păis, Vasile** and Barbu Mititelu, Verginica and Irimia, Elena and Mitrofan, Maria and Gasan, Carol Luca and Micu, Roxana
Titlu: Romanian micro-blogging named entity recognition including health-related entities
Link: <https://aclanthology.org/2022.smm4h-1>

Autori: <u>Păis, Vasile</u> and Mitrofan, Maria and Gasan, Carol Luca and Coneschi, Vlad and Ianov, Alexandru	Publicație: NLLP EMNLP 2021	IF: 0.25	Punctaj: 6
Titlu: Named entity recognition in the Romanian legal domain Link: https://aclanthology.org/2021.nllp-1.2/			
Autori: Avram, Andrei-Marius and <u>Păis, Vasile</u> and Tufiș, Dan	Publicație: NLDB 2021	IF: 0.25	Punctaj: 10
Titlu: A Modular Approach for Romanian-English Speech Translation Link: https://link.springer.com/chapter/10.1007/978-3-030-80599-9_6			
Autori: <u>Păis, Vasile</u> and Mitrofan, Maria	Publicație: SMM4H NAACL 2021	IF: 0.25	Punctaj: 15
Titlu: Assessing multiple word embeddings for named entity recognition of professions and occupations in health-related social media Link: https://aclanthology.org/2021.smm4h-1.27/			
Autori: <u>Păis, Vasile</u> and Ion, Radu	Publicație: CompuTerm LREC 2020	IF: 0.25	Punctaj: 15
Titlu: TermEval 2020: RACAI's automatic term extraction system Link: https://aclanthology.org/2020.computerm-1.14/			
Autori: Tufiș, Dan and Mitrofan, Maria and <u>Păis, Vasile</u> and Ion, Radu and Coman, Andrei	Publicație: LREC 2020	IF: 0.25	Punctaj: 6
Titlu: Collection and Annotation of the Romanian Legal Corpus Link: https://aclanthology.org/2020.lrec-1.337/			
Autori: <u>Păis, Vasile</u> and Tufiș, Dan and Ion, Radu	Publicație: IWLTP LREC 2020	IF: 0.25	Punctaj: 10
Titlu: A Processing Platform Relating Data and Tools for Romanian Language Link: https://aclanthology.org/2020.iwltip-1.13/			
Autori: <u>Păis, Vasile</u> and Tufiș, Dan and Ion, Radu	Publicație: GlobalLex LREC 2020	IF: 0.25	Punctaj: 10

Titlu: RACAI's Word Sense Alignment System using a Similarity Measurement of Dictionary Definitions
Link: <https://aclanthology.org/2020.globalex-1.12/>

Autori: Tamás Váradi, Svetla Koeva, Martin Yamalov, Marko Tadić, Bálint Sass, Bartłomiej Nitoń, Maciej Ogródniczuk, Piotr Pezik and Barbu Mititelu, Verginica and Radu Ion and Irimia, Elena and Mitrofan, Maria and Păis, Vasile and Tufiş, Dan and Radovan Garabík, Simon Krek, Andraz Repar, Matjaž Rihtar and Janez Brank	Publicație: LREC 2020	IF: 0.25	Punctaj: 1.58
Titlu: The MARCELL Legislative Corpus Link: https://aclanthology.org/2020.lrec-1.464/			

Autori: Ion, Radu and Păis, Vasile and Mitrofan, Maria	Publicație: BioNLP EMNLP 2019	IF: 0.25	Punctaj: 10
Titlu: RACAI's System at PharmaCoNER 2019 Link: https://aclanthology.org/D19-5714/			

Autori: Tufiş, Dan and Barbu Mititelu, Verginica and Irimia, Elena and Păis, Vasile and Ion, Radu and Diewald, Nils and Mitrofan, Maria and Mihaela, Onofrei	Publicație: Revue Roumaine de linguistique, 2019	IF: 0.17	Punctaj: 3.55
Titlu: Little Strokes Fell Great Oaks. Creating Corola, The Reference Corpus of Contemporary Romanian Link: http://www.lingv.ro/images/RRL%203%202019%2003-DTufis.pdf			

Autori: Păis, Vasile and Tufiş, Dan	Publicație: CONSILR 2018	IF: 0.25	Punctaj: 15
Titlu: More Romanian word embeddings from the RETEROM project Link: https://profs.info.uaic.ro/~consilr/2019/wp-content/uploads/2019/06/volum-ConsILR-2018-1.pdf#page=99			

Total A2.1. 503.91

A2.2. Articole în reviste și în volumele unor manifestări științifice indexate în alte baze de date internaționale recunoscute (BDI)

Autori: Radu Ion, Verginica Barbu, Vasile Păis , Elena Irimia, Valentin	Publicație: CLIB 2024	Punctaj: 4
---	-------------------------------------	-------------------

Badea				
Titlu: A cross-model study on learning Romanian parts of speech with Transformer models Link: https://dcl.bas.bg/clib/wp-content/uploads/2024/09/CLIB2024_PROCEEDINGS_v1.0.pdf				
Autori: <u>Păis, Vasile</u> and Mîitelu, Verginica Barbu and Ion, Radu and Irimia, Elena	Publicație: SPED 2023			Punctaj: 5
Titlu: Evaluating a Fine-Tuned Whisper Model on Underrepresented Romanian Speech Link: https://ieeexplore.ieee.org/document/10314923				
Autori: Avram, Andrei-Marius and Smădu, Răzvan-Alexandru and <u>Păis, Vasile</u> and Cerceș, Dumitru-Clementin and Ion, Radu and Tușiș, Dan	Publicație: TSP 2023			Punctaj: 3.33
Titlu: Towards Improving the Performance of Pre-Trained Speech Models for Low-Resource Languages Through Lateral Inhibition. Link: https://ieeexplore.ieee.org/document/10197791				
Autori: Barbu, Ana Maria and Irimia, Elena and Vasile, Carmen Mîrzea and <u>Păis, Vasile</u>	Publicație: RANLP 2023			Punctaj: 5
Titlu: Designing the LECOR Learner Corpus for Romanian Link: https://aclanthology.org/2023.ranlp-1.16/				
Autori: Radu Ion, Andrei-Marius Avram, <u>Vasile Păis</u> , Maria Mitrofan, Verginica Barbu Mîitelu, Elena Irimia, Valentin Badea	Publicație: CLIB 2022			Punctaj: 2.86
Titlu: An Open-Domain QA System for e-Governance Link: https://aclanthology.org/2022.clib-1.12/				
Autori: Andrei-Marius Avram, Maria Mitrofan, <u>Vasile Păis</u>	Publicație: IBERLEF 2022			Punctaj: 6.67
Titlu: Species Entity Recognition Using a Neural Inhibitory Mechanism Link: https://dblp.org/rec/conf/sepln/AvramMP22.html				

Autori: <u>Păis, Vasile</u> and Ion, Radu and Avram, Andrei-Marius and Irimia, Elena and Mititelu, Verginica Barbu and Mitrofan, Maria	Publicație: SPED 2021	Punctaj: 3.33
Titlu: Human-Machine Interaction Speech Corpus from the ROBIN project		
Link: https://ieeexplore.ieee.org/document/9587355		

Autori: Avram, Andrei-Marius and <u>Păis, Vasile</u> and Tufiş, Dan	Publicație: RANLP 2021	Punctaj: 6.67
Titlu: PyEuro Voc: A Tool for Multilingual Legal Document Classification with EuroVoc Descriptors		
Link: https://aclanthology.org/2021.ranlp-1.12/		

Autori: Avram, Andrei-Marius and <u>Păis, Vasile</u> and Tufiş, Dan	Publicație: CONSILR 2020	Punctaj: 6.67
Titlu: Romanian speech recognition experiments from the ROBIN project		
Link: https://profs.info.uaic.ro/~consilr/2019/wp-content/uploads/2019/06/volum-ConsILR-2018-1.pdf#page=99		

Autori: <u>Păis, Vasile</u>	Publicație: CONSILR 2020	Punctaj: 20
Titlu: Multiple annotation pipelines inside the RELATE platform		
Link: https://profs.info.uaic.ro/~consilr/2019/wp-content/uploads/2019/06/volum-ConsILR-2018-1.pdf#page=99		

Autori: Barbu Mititelu, Verginica and Irimia, Elena and <u>Păis, Vasile</u> and Avram, Andrei-Marius and Mitrofan, Maria and Cărean, Eric	Publicație: CONSILR 2020	Punctaj: 3.33
Titlu: Romanian Resources in Linguistic Linked Open Data format		
Link: https://profs.info.uaic.ro/~consilr/2019/wp-content/uploads/2019/06/volum-ConsILR-2018-1.pdf#page=99		

Autori: <u>Păis, Vasile</u> and Tufiş, Dan and Ion, Radu	Publicație: CONSILR 2019	Punctaj: 6.67
Titlu: Integration of Romanian NLP tools into the RELATE platform		
Link: https://profs.info.uaic.ro/~consilr/2019/wp-content/uploads/2019/06/volum-ConsILR-2018-1.pdf#page=99		

Autori: Boroş, Tiberiu and Dumitrescu, Ştefan Daniel and Păis, Vasile	Publicație: RoCHI 2018	Punctaj: 6.67
Titlu: Tools and resources for Romanian text-to-speech and speech-to-text applications Link: https://dblp.org/rec/conf/rochi/BorosDP18.html		

Autori: <u>Păis, Vasile</u> and Tufiş, Dan	Publicație: CONSILR 2018	Punctaj: 10
Titlu: More Romanian word embeddings from the RETEROM project Link: https://profs.info.uaic.ro/~consilr/2019/wp-content/uploads/2019/06/volum-ConsILR-2018-1.pdf#page=99		

Total A2.2. 90.2

A2.4. Granturi/proiecte de cercetare

Proiect	Tip	Calitate	Punctaj
System for the anonymisation of Romanian jurisprudence (SAROJ)	International	Director proiect	20
Underrepresented Speech Dataset from Open Data: Case Study on the Romanian Language (USPDATRO)	International	Director proiect	20
ELE	International	Responsabil partener	40
LDS	International	Responsabil partener	40

MARCELL	International	Membru	12
CURLICAT	International	Membru	12
ENRICH4ALL	International	Membru	12
ELG	International	Membru	4
PRESIDENCY	International	Membru	4
ELRC	International	Membru	4
ReTeRom	National	Membru	6

Total A2.4.1.1. Granturi/proiecte de cercetare câștigate prin competiție - Director/responsabil partener - Internationales: 120

Total A2.4.2.1 Granturi/proiecte de cercetare câștigate prin competiție - Membru in echipă - internaționale: 48

Total A2.4.2.2 Granturi/proiecte de cercetare câștigate prin competiție - Membru in echipă - naționale: 6

Total A2. 768.11

A3. Recunoaștere a și impactul activității

Lista completă a citărilor: <https://scholar.google.com/citations?user=kyBZW7UAAA&hl=ro>

Lucrarea citată	Lucrarea care citează	Categoria	Punctaj
Carol-Luca Gasan and Vasile Păiș. 2024. Multi-Model	Ibrahim Said Ahmad, Antonios Anastasopoulos, Ondřej Bojar, Claudia Borg, Marine Carpuat, Roldano Cattani, Mauro Cettolo, William Chen,	ISI	4

<p>System for Effective Subtitling Compression. In Proceedings of the 21st International Conference on Spoken Language Translation (IWSLT 2024), pages 57–64, Bangkok, Thailand (in-person and online). Association for Computational Linguistics.</p>	<p>Qianqian Dong, Marcello Federico, Barry Haddow, Dávid Javorský, Mateusz Krubiński, Tsz Kim Lam, Xutai Ma, Prashant Mathur, Evgeny Matusov, Chandresh Maurya, John McCrae, et al.: 2024. FINDINGS OF THE IWSLT 2024 EVALUATION CAMPAIGN. In Proceedings of the 21st International Conference on Spoken Language Translation (IWSLT 2024), pages 1–11, Bangkok, Thailand (in-person and online). Association for Computational Linguistics.</p>		
<p><u>Păis, V., Mititelu, V. B., Ion, R., & Irimia, E. (2023, October). Evaluating a Fine-Tuned Whisper Model on Underrepresented Romanian Speech. In 2023 International Conference on Speech Technology and Human-Computer Dialogue (SpED) (pp. 141-145). IEEE.</u></p>	<p>Camille Lavigne, Alex Stasica, and Anna Kupsc. 2024. Optimisation des performances d'un système de reconnaissance automatique de la parole pour les commentaires sportifs: fine-tuning de Whisper. In Actes de la 31ème Conférence sur le Traitement Automatique des Langues Naturelles, volume 1 : articles longs et prises de position, pages 567–581, Toulouse, France. ATALA and AFPC.</p>	<p>BDI</p>	<p>1</p>
<p>Barbu, A. M., Irimia, E., Vasile, C. M., & <u>Păis, V. (2023, September). Designing the LECOR Learner Corpus for Romanian. In Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing (pp. 143-152).</u></p>	<p>Niculescu, O. (2023). Modal versus creaky filler particles in Romanian connected speech. <i>Revue Roumaine de Linguistique</i>, 68, 3-4.</p>	<p>ISI</p>	<p>2</p>

<p>Avram, A. M., Mititelu, V. B., Păis, V., Cercel, D. C., & Trăușan-Matu, Ș. (2023). Multilingual Multiword Expression Identification and Using Lateral Inhibition and Domain Adaptation. <i>Mathematics</i>, 11(11), 2548.</p>	<p>Lal, M., Bhende, M., Sharma, S., Thorat, P., Goel, A., Tamrakar, P., & Pathak, A. (2023, December). Next-Gen Language Mastery: Exploring Advances in Natural Language Processing Post-transformers. In <i>International Conference on Cognitive Computing and Cyber Physical Systems</i> (pp. 453-469). Singapore: Springer Nature Singapore.</p>	<p>ISI</p>	<p>1.6</p>
<p>Păis, Vasile. RACAI at SemEval-2022 Task 11: Complex named entity recognition using a lateral inhibition mechanism. In <i>Proceedings of the 16th International Workshop on Semantic Evaluation</i> (SemEval-2022), Association for Computational Linguistics, pp. 1562-1569, SemEval NAACL, 2022</p>	<p>Shervin Malmasi, Anjie Fang, Besnik Fetahu, Sudipta Kar, and Oleg Rokhlenko. 2022. SemEval-2022 Task 11: Multilingual Complex Named Entity Recognition (MultiCoNER). In <i>Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)</i>, pages 1412-1437, Seattle, United States. Association for Computational Linguistics.</p>	<p>ISI</p>	<p>8</p>
<p>Proceedings of the 16th International Workshop on Semantic Evaluation</p>	<p>Parsaimehr, E., Fartash, M. & Akbari Torkestani, J. Improving Feature Extraction Using a Hybrid of CNN and LSTM for Entity Identification. <i>Neural Process Lett</i> (2023).</p>	<p>ISI</p>	<p>8</p>
<p>(SemEval-2022), Association for Computational Linguistics, pp. 1562-1569, SemEval NAACL, 2022</p>	<p>Wang Y, Zhao Q, Ma M, Xu J. Decoding Structure-Odor Relationship Based on Hypergraph Neural Network and Deep Attentional Factorization Machine. <i>Applied Sciences</i>. 2022; 12(17):8777.</p>	<p>BDI</p>	<p>4</p>
<p>SemEval NAACL, 2022</p>	<p>Iker García-Ferrero, Jon Ander Campos, Oscar Sainz, Ander Salaberria, and Dan Roth. 2023. IXA/Cogcomp at SemEval-2023 Task 2: Context-enriched Multilingual Named Entity Recognition Using Knowledge Bases. In <i>Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)</i>, pages 1335-1346, Toronto, Canada. Association for Computational Linguistics.</p>	<p>BDI</p>	<p>4</p>
<p>Păis, V., Mititelu, V. B., Irimia, E., Mitrofan, M.,</p>	<p>Mănescu, E. C., Smădu, R. A., Avram, A. M., Cercel, D. C., & Pop, F. (2023, October). End-to-End Lip Reading in Romanian with Cross-Lingual Domain Adaptation and Lateral Inhibition. In <i>2023 IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)</i> (pp. 287-293). IEEE.</p>	<p>BDI</p>	<p>4</p>
<p></p>	<p>Tăiatu, I. M., Avram, A. M., Cercel, D. C., & Pop, F. (2023, June). RoBERTweet: A BERT Language Model for Romanian Tweets. In</p>	<p>ISI</p>	<p>1.33</p>

<p>Gasan, C. L., & Micu, R. (2022, October). Romanian micro-blogging named entity recognition including health-related entities. In Proceedings of The Seventh Workshop on Social Media Mining for Health Applications, Workshop & Shared Task (pp. 190-196).</p>	<p>International Conference on Applications of Natural Language to Information Systems (pp. 577-587). Cham: Springer Nature Switzerland.</p>		
<p>Maria Mitrofan and Vasile Păiș. 2022. Improving Romanian BioNER Using a Biologically Inspired System. In Proceedings of the 21st Workshop on Biomedical Language Processing, pages 316–322, Dublin, Ireland. Association for Computational Linguistics.</p>	<p>Cuéllar Hidalgo, R., Pinto Elías, R., Torres-Moreno, J. M., Vergara Villegas, O. O., Reyes Salgado, G., & Magadán Salazar, A. (2024). Neural Architecture Comparison for Bibliographic Reference Segmentation: An Empirical Study. <i>Data</i>, 9(5), 71.</p>	<p>ISI</p>	<p>4</p>
	<p>Lupancu, V. C., Platica, A. G., Rosu, C. M., Gifu, D., & Trandabat, D. (2023, July). FII_Better at SemEval-2023 Task 2: MultiCoNER II Multilingual Complex Named Entity Recognition. In Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023) (pp. 1107-1113).</p>	<p>ISI</p>	<p>4</p>
	<p>Lupancu, V. C., & Iftene, A. (2023). Multilingual Fine-Grained Named Entity Recognition. <i>Computer Science Journal of Moldova</i>, 93(3), 321-339.</p>	<p>BDI</p>	<p>2</p>
<p>Radu Ion, Andrei-Marius Avram, Vasile Păiș, Maria Mitrofan, Verginica Barbu</p>	<p>Irimia, E., & Mititelu, V. (2023, January). The Romanian Wordnet in Linked Open Data Format. In Proceedings of the 12th Global Wordnet Conference (pp. 127-132).</p>	<p>BDI</p>	<p>0.57</p>
<p>Mititelu, Elena Irimia, and Valentin Badea. 2022. An Open-Domain QA System for e-Governance. In Proceedings of the 5th International Conference on Computational Linguistics in Bulgaria (CLIB 2022), pages 105–112, Sofia,</p>	<p>Gu, S., Wang, J., Zheng, S., Pan, Y., Jiang, G., Dai, W., ... & Chen, D. (2022, November). Knowledge Graph Based Chicken Disease Diagnosis Question Answering System. In International Conference on Data Mining and Big Data (pp. 246-259). Singapore: Springer Nature Singapore.</p>	<p>BDI</p>	<p>0.57</p>

<p>Bulgaria. Department of Computational Linguistics, IBL -- BAS.</p> <p><u>Păis, Vasile</u> and Ion, Radu and Avram, Andrei-Marius and Irimia, Elena and Mititelu, Verginica Barbu and Mitrofan, Maria. Human-Machine Interaction Speech Corpus from the ROBIN project. In 2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD). vol. , no. , pp. 91-96, 2021</p>	<p>Tufiș, D. (2022). Romanian Language Technology – a view from an academic perspective. <i>INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL</i>, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	<p>ISI</p>	<p>1.33</p>
<p>Avram, Andrei-Marius and <u>Păis, Vasile</u> and Mitrofan, Maria. 2022. RACAI@SMM4H'22: Tweets Disease Mention Detection Using a Neural Lateral Inhibitory Mechanism. In Proceedings of The Seventh Workshop on Social Media Mining for Health Applications, Workshop & Shared Task, pages 1–3, Gyeongju, Republic of Korea. Association for Computational Linguistics.</p>	<p>Davy Weissenbacher, Juan Banda, Vera Davydova, Darryl Estrada Zavala, Luis Gasco Sánchez, Yao Ge, Yuting Guo, Ari Klein, Martin Krallinger, Mathias Leddin, Arjun Magge, Raul Rodriguez-Esteban, Abeed Sarker, Lucia Schmidt, Elena Tutubalina, and Graciela Gonzalez-Hernandez. 2022. Overview of the Seventh Social Media Mining for Health Applications (#SMM4H) Shared Tasks at COLING 2022. In Proceedings of The Seventh Workshop on Social Media Mining for Health Applications, Workshop & Shared Task, pages 221–241, Gyeongju, Republic of Korea. Association for Computational Linguistics.</p>	<p>ISI</p>	<p>2.67</p>
<p>Luis Gasco Sánchez, Darryl Estrada Zavala, Eulàlia Farré-Maduell, Salvador Lima-López, Antonio Miranda-Escalada, and Martin Krallinger. 2022. The SocialDisNER shared task on detection of disease mentions in health-relevant content from social media: methods, evaluation, guidelines and corpora. In Proceedings of The Seventh Workshop on Social Media Mining for Health Applications, Workshop & Shared Task, pages 182–189, Gyeongju, Republic of Korea. Association for Computational Linguistics.</p> <p>Mănescu, E. C., Smădu, R. A., Avram, A. M., Cercel, D. C., & Pop, F.</p>	<p>Luis Gasco Sánchez, Darryl Estrada Zavala, Eulàlia Farré-Maduell, Salvador Lima-López, Antonio Miranda-Escalada, and Martin Krallinger. 2022. The SocialDisNER shared task on detection of disease mentions in health-relevant content from social media: methods, evaluation, guidelines and corpora. In Proceedings of The Seventh Workshop on Social Media Mining for Health Applications, Workshop & Shared Task, pages 182–189, Gyeongju, Republic of Korea. Association for Computational Linguistics.</p>	<p>ISI</p>	<p>2.67</p>
		<p>BDI</p>	<p>1.33</p>

	<p>(2023, October). End-to-End Lip Reading in Romanian with Cross-Lingual Domain Adaptation and Lateral Inhibition. In 2023 IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) (pp. 287-293). IEEE.</p>		
<p>Váradí, Tamás and Nyéki, Bence and Koeva, Svetla and Tadić, Marko and Štefanec, Vanja and Ogrodniczuk, Maciej and Nitoń, Bartłomiej and Pezik, Piotr and Barbu Mititelu, Verginica and Irimia, Elena and Mitrofan, Maria and <u>Păis, Vasile</u> and Tufiş, Dan and Garabîk, Radovan and Krek, Simon and Repar, Andraş. 2022. Introducing the CURLICAT Corpora: Seven-language Domain Specific Annotated Corpora from Curated Sources. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pages 100–108, Marseille, France. European Language Resources Association.</p>	<p>Szczerbicki, E., Wojtkiewicz, K., Van Nguyen, S., Pietranik, M., & Krótkiewicz, M. (Eds.). (2022). Recent Challenges in Intelligent Information and Database Systems: 14th Asian Conference, ACIIDS 2022, Ho Chi Minh City, Vietnam, November 28-30, 2022, Proceedings. Springer Nature.</p> <p>Ogrodniczuk, Maciej. "Fine-Tuning OCR Error Detection and Correction in a Polish Corpus of Scientific Abstracts." Recent Challenges in Intelligent Information and Database Systems: 14th Asian Conference, ACIIDS 2022, Ho Chi Minh City, Vietnam, November 28-30, 2022, Proceedings. Singapore: Springer Nature Singapore, 2022.</p>	<p>ISI</p>	<p>0.5</p>
<p><u>Păis, Vasile</u> and Ion, Radu and Avram, Andrei-Marius and Mitrofan, Maria and Tufiş, Dan. In-depth evaluation of Romanian</p>	<p>Chai, C. (2022). Comparison of text preprocessing methods. Natural Language Engineering, 1-45. doi:10.1017/S1351324922000213</p> <p>Avram, A. M., Catrina, D., Cercel, D. C., Dascalu, M., Rebedea, T., Păiş, V., & Tufiş, D. (2022, June). Distilling the Knowledge of Romanian BERTs Using Multiple Teachers. In Proceedings of the Thirteenth Language</p>	<p>ISI</p>	<p>1.6</p>
		<p>ISI</p>	<p>1.6</p>

natural language processing pipelines. In Romanian Journal of Information Science and Technology (ROMJIST). vol. 24, no. 4, pp. 384-401, 2021	Resources and Evaluation Conference (pp. 374-384). Bucur, A. M., Dincă, A., Chitez, M., & Rogobete, R. (2023, September). Automatic Extraction of the Romanian Academic Word List: Data and Methods. In Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing (pp. 234-241).	BDI	0.8
Avram, A. M., <u>Păis, V.</u> , & Tufis, D. I. (2021, September). PyEuroVoc: A Tool for Multilingual Legal Document Classification with EuroVoc Descriptors. In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021) (pp. 92-101).	Linders, G. M., & Louwerse, M. M. (2023). Lingualyzer: A computational linguistic tool for multilingual and multidimensional text analysis. Behavior research methods, 1-28.	ISI	1.6
	Greco, C. M., & Tagarelli, A. (2023). Bringing order into the realm of Transformer-based language models for artificial intelligence and law. Artificial Intelligence and Law, 1-148.	ISI	2.67
	Gupta, S., Gopi, G., Balaji, H., Poplavska, E., O'Toole, N., Arora, S., ... & Wilson, S. (2024, May). Creation and Analysis of an International Corpus of Privacy Laws. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024) (pp. 4092-4105).	ISI	2.67
	Rovera, M., Aprosio, A. P., Greco, F., Lucchese, M., Tonelli, S., & Antetomaso, A. (2023, December). Italian Legislative Text Classification for Gazzetta Ufficiale. In Proceedings of the Natural Legal Language Processing Workshop 2023 (pp. 44-50).	ISI	2.67
	Trancoso, I., Mamede, N., Martins, B., Pinto, H. S., & Ribeiro, R. (2023). The Impact of Language Technologies in the Legal Domain. In Multidisciplinary Perspectives on Artificial Intelligence and the Law (pp. 25-46). Cham: Springer International Publishing.	BDI	1.33
	Hyvönen, E., Cao, H., Leal, R., Rantala, H., & Hietanen, A. (2024). A Model and Case Study for Searching and Reading Cross-Border Multilingual Legislation on the Semantic Web. In Knowledge Graphs in the Age of Language Models and Neuro-Symbolic AI (pp. 261-278). IOS Press.	Carte	2.67
	Qin, R., Huang, M., & Luo, Y. (2022, May). A comparison study of pre-trained language models for Chinese legal document classification. In 2022 5th International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 444-449). IEEE.	BDI	1.33

<p>Păis, Vasile, Maria Mitrofan, Carol Luca Gasan, Vlad Coneschi, Alexandru Ianov, Named entity recognition in the Romanian legal domain, Proceedings of the Natural Language Processing Workshop 2021, pp 9-18, NLLP EMNLP 2021</p>	<p>Yuquan Le, Sheng Xiao, Zheng Xiao, Kenli Li, Topology-aware Multi-task Learning Framework for Civil Case Judgment Prediction, Expert Systems with Applications, Volume 238, Part F, 2024, 122103, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2023.122103.</p>	<p>ISI</p>	<p>1.6</p>
<p>Greco, C.M., Tagarelli, A. Bringing order into the realm of Transformer-based language models for artificial intelligence and law. Artif Intell Law (2023). https://doi.org/10.1007/s10506-023-09374-7</p>	<p>X. Zhang, X. Luo and J. Wu, "A RoBERTa-GlobalPointer-Based Method for Named Entity Recognition of Legal Documents," 2023 International Joint Conference on Neural Networks (IJCNN), Gold Coast, Australia, 2023, pp. 1-8, doi: 10.1109/IJCNN54540.2023.10191275.</p>	<p>ISI</p>	<p>1.6</p>
<p>Luis Adrián Cabrera-Diego and Akshita Gheewala. 2023. Jus Mundi at SemEval-2023 Task 6: Using a Frustratingly Easy Domain Adaption for a Legal Named Entity Recognition System. In Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023), pages 1783–1790, Toronto, Canada. Association for Computational Linguistics.</p>	<p>Echim, SV, Smădu, RA., Avram, AM., Cercel, DC., Pop, F. (2023). Adversarial Capsule Networks for Romanian Satire Detection and Sentiment Analysis. In: Métais, E., Meziane, F., Sugumaran, V., Manning, W., Reiff-Marganec, S. (eds) Natural Language Processing and Information Systems. NLDB 2023. Lecture Notes in Computer Science, vol 13913. Springer, Cham. https://doi.org/10.1007/978-3-031-35320-8_31</p>	<p>BDI</p>	<p>0.8</p>
<p>Niklaus, J., Matoshi, V., Rani, P., Galassi, A., Stürmer, M., & Chalkidis, I. (2023, December). LEXTREME: A Multi-Lingual and Multi-Task Benchmark for the Legal Domain. In Findings of the Association for Computational Linguistics: EMNLP 2023 (pp. 3016-3054).</p>	<p>Kalamkar, P., Agarwal, A., Tiwari, A., Gupta, S., Kam, S., & Raghavan, V. (2022). Named Entity Recognition in Indian court judgments. NLLP 2022, 184-193.</p>	<p>ISI</p>	<p>1.6</p>
<p>Au, T. W. T., Lampos, V., & Cox, I. (2022, December). E-NER—An Annotated Named Entity Recognition Corpus of Legal Text. In Proceedings of the Natural Language Processing Workshop 2022 (pp. 246-255).</p>	<p></p>	<p>ISI</p>	<p>1.6</p>

	<p>Cabrera-Diego, L. A., & Gheewala, A. (2023, July). Jus mundi at SemEval-2023 task 6: Using a frustratingly easy domain adaption for a legal named entity recognition system. In Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023) (pp. 1783-1790).</p>	<p>ISI</p>	<p>1.6</p>
	<p>Claire Barale, Michael Rovatsos, and Nehal Bhuta. 2023. Automated Refugee Case Analysis: A NLP Pipeline for Supporting Legal Practitioners. In Findings of the Association for Computational Linguistics: ACL 2023, pages 2992–3005, Toronto, Canada. Association for Computational Linguistics.</p>	<p>BDI</p>	<p>0.8</p>
	<p>Tran, Quang-Minh, and Xuan-Dung Doan. "VTCC-NER at SemEval-2023 Task 6: An Ensemble Pre-trained Language Models for Named Entity Recognition." Proceedings of the The 17th International Workshop on Semantic Evaluation (SemEval-2023). 2023.</p>	<p>BDI</p>	<p>0.8</p>
	<p>Tuفیş, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	<p>ISI</p>	<p>1.6</p>
	<p>Bernsohn, D., Semo, G., Vazana, Y., Hayat, G., Hagag, B., Niklaus, J., ... & Truskovskiy, K. (2024, March). Legallens: Leveraging LLMs for Legal Violation Identification in Unstructured Text. In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 2129-2145).</p>	<p>ISI</p>	<p>1.6</p>
	<p>Zhang, Y., & Xiao, G. (2024). Named Entity Recognition Datasets: A Classification Framework. International Journal of Computational Intelligence Systems, 17(1), 71.</p>	<p>ISI</p>	<p>1.6</p>
	<p>Gupta, S., Gopi, G., Balaji, H., Poplavska, E., O'Toole, N., Arora, S., ... & Wilson, S. (2024, May). Creation and Analysis of an International Corpus of Privacy Laws. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024) (pp. 4092-4105).</p>	<p>ISI</p>	<p>1.6</p>
	<p>Bucur, A. M., Dincă, A., Chitez, M., & Rogobete, R. (2023, September). Automatic Extraction of the Romanian Academic Word List: Data and</p>	<p>BDI</p>	<p>0.8</p>

<p>Paiş, Vasile and Tufiş, Dan. Capitalization and punctuation restoration: a survey. In <i>Artificial Intelligence Review</i>, vol 55, issue 3, pp 1681-1722, 2022</p>	<p>Methods. In <i>Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing</i> (pp. 234-241).</p> <p>Nunes, R. O., Balreira, D. G., Spritzer, A. S., & Freitas, C. M. D. S. (2024, March). A Named Entity Recognition Approach for Portuguese Legislative Texts Using Self-Learning. In <i>Proceedings of the 16th International Conference on Computational Processing of Portuguese</i> (pp. 290-300).</p> <p>Smădu, R. A., Dinićă, I. R., Avram, A. M., Cercel, D. C., Pop, F., & Cercel, M. C. (2022, December). Legal named entity recognition with multi-task domain adaptation. In <i>Proceedings of the Natural Legal Language Processing Workshop 2022</i> (pp. 305-321).</p> <p>Lee, S. M., Tan, Y. H., & Yu, H. T. (2023, June). LeArNER: few-shot legal argument named entity recognition. In <i>Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law</i> (pp. 422-426).</p> <p>Corina Ceausu and Sergiu Nisioi. 2022. Identifying Draft Bills Impacting Existing Legislation: a Case Study on Romanian. In <i>Proceedings of the Thirteenth Language Resources and Evaluation Conference</i>, pages 3670–3674, Marseille, France. European Language Resources Association.</p> <p>Islam, T., Kundu, A., Lima, R. J., Hena, M. H., Sharif, O., Rahman, A., & Hasan, M. Z. (2023). 5 Review Analysis of Ride-Sharing Applications Using Machine Learning Approaches. <i>Computational Statistical Methodologies and Modeling for Artificial Intelligence</i>, 99-122.</p> <p>Vandeghinste, V., Guhr, O. FullStop: punctuation and segmentation prediction for Dutch with transformers. <i>Lang Resources & Evaluation</i> (2023). https://doi.org/10.1007/s10579-023-09676-x</p> <p>Bakare, A. M., Anbananthen, K. S. M., Muthaiyah, S., Krishnan, J., & Kannan, S. (2023). Punctuation restoration with transformer model on social media data. <i>Applied Sciences</i>, 13(3), 1685.</p>	<p>BDI</p> <p>ISI</p> <p>ISI</p> <p>ISI</p> <p>ISI</p> <p>Carte</p> <p>ISI</p> <p>ISI</p>	<p>0.8</p> <p>1.6</p> <p>1.6</p> <p>1.6</p> <p>4</p> <p>4</p> <p>4</p>
---	---	---	---

	<p>Sheik, R., Ganta, S. R., & Nirmala, S. J. (2024). Legal sentence boundary detection using hybrid deep learning and statistical models. <i>Artificial Intelligence and Law</i>, 1-31.</p> <p>Minixhofer, B., Pfeiffer, J., & Vulić, I. (2023, July). Where's the Point? Self-Supervised Multilingual Punctuation-Agnostic Sentence Segmentation. In <i>Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)</i> (pp. 7215-7235).</p> <p>Piyush Behre, Sharman Tan, Padma Varadharajan and Shuangyu Chang. 2022. Streaming punctuation: A novel punctuation technique leveraging bidirectional context for continuous speech recognition. <i>International Journal on Natural Language Computing (IJNLC) Vol.11, No.6, December 2022</i></p> <p>De Lima, T.B. et al. (2022). Sequence Labeling Algorithms for Punctuation Restoration in Brazilian Portuguese Texts. In: Xavier-Junior, J.C., Rios, R.A. (eds) <i>Intelligent Systems. BRACIS 2022. Lecture Notes in Computer Science()</i>, vol 13654 . Springer, Cham.</p> <p>De Lima, T. B., Rolim, V., Nascimento, A. C., Miranda, P., Macario, V., Rodrigues, L., ... & Mello, R. F. (2024). Towards explainable automatic punctuation restoration for Portuguese using transformers. <i>Expert Systems with Applications</i>, 257, 125097.</p> <p>Sakr, A., & Torki, M. (2023, December). AraPunc: Arabic Punctuation Restoration Using Transformers. In <i>2023 20th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)</i> (pp. 1-6). IEEE.</p> <p>Kuhn, K., Kersken, V., Zimmermann, G. (2024) Beyond Levenshtein: Leveraging Multiple Algorithms for Robust Word Error Rate Computations And Granular Error Classifications. <i>Proc. Interspeech 2024</i>, 4543-4547, doi: 10.21437/Interspeech.2024-32</p> <p>Meister, A., Novikov, M., Karpov, N., Bakhturina, E., Lavrukhin, V., &</p>	<p>ISI</p> <p>ISI</p> <p>BDI</p> <p>ISI</p> <p>ISI</p> <p>ISI</p> <p>ISI</p> <p>BDI</p>	<p>4</p> <p>4</p> <p>2</p> <p>4</p> <p>4</p> <p>4</p> <p>4</p> <p>2</p>
--	---	---	---

<p>Păis, Vasile and Mitrofan, Maria. Assessing multiple word embeddings for named entity recognition of professions and occupations in health-related social media. In Proceedings of the Sixth Social Media Mining for Health (#SMM4H) Workshop and Shared Task. Association for Computational Linguistics, Mexico City, Mexico, pp. 128--130, jun 2021</p>	<p>Ginsburg, B. (2023, December). LibriSpeech-PC: Benchmark for Evaluation of Punctuation and Capitalization Capabilities of end-to-end ASR Models. In 2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU) (pp. 1-7). IEEE.</p> <p>Yagi, S., Fareh, S., Elnagar, A., Balajeed, M., El-mneizel, A., & Al-Badawi, M. (2024). Is Arabic punctuation rule-governed?. Cogent Arts & Humanities, 11(1), 2303818.</p> <p>Machura, J., Frémund, A., Švec, J. (2022). Automatic Grammar Correction of Commas in Czech Written Texts: Comparative Study. In: Sojka, P., Horák, A., Kopeček, I., Pala, K. (eds) Text, Speech, and Dialogue. TSD 2022. Lecture Notes in Computer Science(), vol 13502. Springer, Cham. https://doi.org/10.1007/978-3-031-16270-1_10</p> <p>MUNFADLILA, A. W. (2022). Analysis of The Use of Punctuation Marks on The Writing Ability of Nursing Students. Journal of English Development, 2(02), 57–74.</p> <p>Arjun Magge, Ari Klein, Antonio Miranda-Escalada, Mohammed Ali Al-Garadi, Ilseyar Alimova, Zulfat Miftahudinov, Eulalia Farre, Salvador Lima López, Ivan Flores, Karen O’Connor, Davy Weissenbacher, Elena Tutubalina, Abeer Sarker, Juan Banda, Martin Krallinger, Graciela Gonzalez-Hernandez, Overview of the Sixth Social Media Mining for Health Applications (#SMM4H) Shared Tasks at NAACL 2021, In Proceedings of the Sixth Social Media Mining for Health (#SMM4H) Workshop and Shared Task, pp 21-32, NAACL, 2021.</p> <p>Antonio Miranda-Escalada, Eulàlia Farré-Maduell, Salvador Lima-López, Luis Gascó, Vicent Briva-Iglesias, Marvín Agüero-Torales, Martín Krallinger, The ProfNER shared task on automatic recognition of occupation mentions in social media: systems, evaluation, guidelines, embeddings and corpora, In Proceedings of the Sixth Social Media Mining for Health (#SMM4H) Workshop and Shared Task, pp 13–20, NAACL, 2021.</p>	<p>BDI</p> <p>ISI</p> <p>BDI</p> <p>ISI</p> <p>BDI</p> <p>ISI</p> <p>BDI</p> <p>ISI</p>	<p>2</p> <p>4</p> <p>2</p> <p>4</p> <p>4</p> <p>4</p>
---	--	---	---

<p><u>Păis, Vasile</u> and Ion, Radu. TermEval 2020: RACAI's automatic term extraction system. In Proceedings of the 6th International Workshop on Computational Terminology. European Language Resources Association, Marseille, France, pp. 101-105, May 2020</p>	<p>Rigouts Terryn, A., Hoste, V., Drouin, P., & Lefever, E. (2020). TermEval 2020 : shared task on automatic term extraction using the Annotated Corpora for term Extraction Research (ACTER) dataset. In B. Daile, K. Kageura, & A. Rigouts Terryn (Eds.), 6th International Workshop on Computational Terminology (COMPUTERM 2020), Proceedings (pp. 85–94). Marsaille, france: European Language Resources Association (ELRA).</p> <p>Kafando R, Decoupes R, Valentin S, Sautot L, Teisseire M, Roche M. ITEXT-BIO: Intelligent Term EXtraction for BIOMedical analysis. Health Inf Sci Syst. 2021 Jul 10;9(1):29. doi: 10.1007/s13755-021-00156-6. PMID: 34276970; PMCID: PMC8272612.</p> <p>Aliya Nugumanova, Darkhan Akhmed-Zaki, Madina Mansurova, Yerzhan Baiburin, Almasbek Maulit (2022). NMF-based approach to automatic term extraction, Expert Systems with Applications, Volume 199, 117179, ISSN 0957-4174.</p>	<p>BDI</p>	<p>2</p>
<p>Tufiş, Dan and Mitrofan, Maria and <u>Păis, Vasile</u> and Ion, Radu and Coman, Andrei. Collection and Annotation of the Romanian Legal Corpus. In Proceedings of The 12th Language Resources and Evaluation Conference. European Language Resources Association, Marseille, France, pp. 2766-2770, May 2020</p>	<p>PETRU REBEJA, EXPLORING VARIATIONAL AUTOENCODERS FOR LEMMATIZATION, CONSILR 2020, pp 77-82</p>	<p>BDI</p>	<p>0.8</p>
<p><u>Păis, Vasile</u> and Tufiş, Dan and Ion, Radu. A Processing Platform Relating Data and Tools for Romanian Language. In Proceedings of</p>	<p>Tufiş, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	<p>ISI</p>	<p>1.6</p>
<p><u>Păis, Vasile</u> and Tufiş, Dan and Ion, Radu. A Processing Platform Relating Data and Tools for Romanian Language. In Proceedings of</p>	<p>Bulgaru, Gabriela. "Linguistic resources for Romanian. Quality assessed machine-translated and post-edited Romanian corpus and opportunities for interdisciplinary research." Revue Internationale d'Études en Langues Modernes Appliquées 14.Suppl. (2021): 78-95.</p>	<p>BDI</p>	<p>1.33</p>

<p>The 12th Language Resources and Evaluation Conference. European Language Association, Marseille, France, pp. 81-88, May 2020</p>	<p>Balauta, Ioana Daniela. "Corpus de ghiduri turistice generale despre România, redactate în română și franceză. Constituire, tratare, instrumente de lucru și tipuri de analize." Meridian Critic-Analele Universității Ștefan cel Mare" Suceava 38 (2022): 319-337.</p>	<p>BDI</p>	<p>1.33</p>
<p><u>Păis, Vasile</u> and Tufiș, Dan and Ion, Radu. MWSA Task at GlobalLex 2020: RACAI's Word Sense Alignment System using a Similarity Measurement of Dictionary Definitions. In Proceedings of The 12th Language Resources and Evaluation Conference. European Language Association, Marseille, France, pp. 69-75, May 2020</p>	<p>Mititelu, V. B., Cristescu, M., Mitrofan, M., Zgreabăn, B. M., & Bărbulescu, E. A. (2022, September). A Romanian Treebank Annotated with Verbal Multiword Expressions. In Proceedings of the 5th International Conference on Computational Linguistics in Bulgaria (CLIB 2022) (pp. 137-145).</p>	<p>BDI</p>	<p>1.33</p>
<p>Tufiș, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	<p>Tufiș, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	<p>ISI</p>	<p>2.67</p>
<p><u>Păis, Vasile</u> and Tufiș, Dan and Ion, Radu. MWSA Task at GlobalLex 2020: RACAI's Word Sense Alignment System using a Similarity Measurement of Dictionary Definitions. In Proceedings of The 12th Language Resources and Evaluation Conference. European Language Association, Marseille, France, pp. 69-75, May 2020</p>	<p>Hagen, Thora, Fotis Jannidis, and Andreas Witt. "Word sense alignment and disambiguation for historical encyclopedias." Graphs and Networks in the Humanities 2022. Technologies, Models, Analyses, and Visualizations. 6th International Conference, 3.-4. February 2022, Online.. Graphen & Netzwerke; AG des Verbandes Digital Humanities im deutschsprachigen Raum eV, 2022.</p>	<p>BDI</p>	<p>1.33</p>

<p>Avram, Andrei-Marius and Păiș, Vasile and Tufiş, Dan. Towards a Romanian end-to-end automatic speech recognition based on Deepspeech2. In Proceedings of the Romanian Academy Series A. vol. 21, pp. 395--402, 2020</p>	<p>Albuquerque, H.O. et al. (2022). UlyssesNER-Br: A Corpus of Brazilian Legislative Documents for Named Entity Recognition. In: Computational Processing of the Portuguese Language. PROPOR 2022. Lecture Notes in Computer Science(), vol 13208. Springer, Cham. https://doi.org/10.1007/978-3-030-98305-5_1</p> <p>Tamás Váradi, Marko Tadić, Svetla Koeva, Maciej Ogródniczuk, Dan Tufiş, Radovan Garabík, Simon Krek, and Andraž Repar. 2022. Curated Multilingual Language Resources for CEF AT (CURLICAT): overall view. In Proceedings of the 23rd Annual Conference of the European Association for Machine Translation, pages 339–340, Ghent, Belgium. European Association for Machine Translation.</p> <p>Corina Ceausu and Sergiu Nisioi. 2022. Identifying Draft Bills Impacting Existing Legislation: a Case Study on Romanian. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pages 3670–3674, Marseille, France. European Language Resources Association.</p> <p>Márton Makrai, Ákos Máté Tündik, Balázs Indig, György Szaszák. Towards abstractive summarization in Hungarian. In Proceedings of Magyar Számítógépes Nyelvészeti Konferencia, pp 505-519, Szeged, 2022.</p> <p>Zgank A. IoT-Based Bee Swarm Activity Acoustic Classification Using Deep Neural Networks. Sensors. 2021; 21(3):676. https://doi.org/10.3390/s21030676</p> <p>George-Eduard Zaharia, Andrei-Marius Avram, Dumitru-Clementin Cercel, Traian Rebedea, Dialect Identification through Adversarial Learning and Knowledge Distillation on Romanian BERT, In Proceedings of the Eighth Workshop on NLP for Similar Languages, Varieties and Dialects, EACL, pp 113-119, 2021.</p> <p>Tufiş, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	<p>ISI</p> <p>ISI</p> <p>BDI</p> <p>BDI</p> <p>ISI</p> <p>ISI</p> <p>ISI</p>	<p>0.42</p> <p>0.42</p> <p>0.21</p> <p>0.21</p> <p>2.67</p> <p>2.67</p> <p>2.67</p>
---	---	--	---

<p>Avram, Andrei-Marius and Păis, Vasile and Tufiş, Dan. Romanian speech recognition experiments from the ROBIN project. In The 15th International Conference on Linguistic Resources and Tools for Natural Language Processing. pp. 103--114, 2020</p>	<p>D. Ungureanu, M. Badeanu, G. -C. Marica, M. Dascalu and D. I. Tufiş, "Establishing a Baseline of Romanian Speech-to-Text Models," 2021 International Conference on Speech Technology and Human-Computer Dialogue (SpED), 2021, pp. 132-138, doi: 10.1109/SpED53181.2021.9587345.</p>	<p>BDI</p>	<p>1.33</p>
<p>Păis, Vasile. Multiple annotation pipelines inside the RELATE platform. In The 15th International Conference on Linguistic Resources and Tools for Natural Language Processing. pp. 65--75, 2020</p>	<p>Tufiş, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	<p>ISI</p>	<p>2.67</p>
<p>Păis, Vasile. Multiple annotation pipelines inside the RELATE platform. In The 15th International Conference on Linguistic Resources and Tools for Natural Language Processing. pp. 65--75, 2020</p>	<p>A. -M. Avram, M. -V. Nichita, R. -G. Bartusica and M. -V. Mihai, "RoSAC: A Speech Corpus for Transcribing Romanian Emergency Calls," 2022 14th International Conference on Communications (COMM), Bucharest, Romania, 2022, pp. 1-5, doi: 10.1109/COMM54429.2022.9817214.</p>	<p>BDI</p>	<p>1.33</p>
<p>Păis, Vasile. Multiple annotation pipelines inside the RELATE platform. In The 15th International Conference on Linguistic Resources and Tools for Natural Language Processing. pp. 65--75, 2020</p>	<p>Tufiş, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	<p>ISI</p>	<p>8</p>
<p>Ion, Radu and Păis, Vasile and Mitrofan, Maria. RACAI's System at PharmaCoNER 2019. In In Proceedings of the 5th Workshop on BioNLP Open Shared Tasks. (c) 2019 Association for Computational Linguistics, pp. 90-99, 2019</p>	<p>Mititelu, V.B. and Mitrofan, M., The Romanian Medical Treebank - SiMoNERo. Proceedings of the The 15th Edition of the International Conference on Linguistic Resources and Tools for Natural Language Processing – ConSLLR-2020 ISSN 1843-911X, p.7-16, 2020.</p>	<p>BDI</p>	<p>1.33</p>

<p>Academy Series A - Mathematics Physics Technical Sciences Information Science. vol. 19, no. 2, pp. 185--191, 2018</p>	<p>57, 2019, Pages 128-145, ISSN 0885-2308, https://doi.org/10.1016/j.csl.2019.02.005. D. Tufiş, V. Barbu Mititelu, E. Irimia, M. Mitrofan, R. Ion and G. Cioroiu, "Making Pepper Understand and Respond in Romanian," 2019 22nd International Conference on Control Systems and Computer Science (CSCS), 2019, pp. 682-688, doi: 10.1109/CSCS.2019.00122.</p>		
	<p>Dan Tufiş, Dan Cristea, A Bird's-eye View of Language Processing Projects at the Romanian Academy, LREC 2018, L18-1388, 2018</p>	ISI	4
	<p>Elena Irimia, Maria Mitrofan, Verginica Mititelu, Evaluating the Wordnet and CoRoLa-based Word Embedding Vectors for Romanian as Resources in the Task of Microworlds Lexicon Expansion, Proceedings of the 10th Global Wordnet Conference, pp 176-184, 2019</p>	BDI	2
	<p>Raicu I., Boitout N., Bologa R., Sturza M.G., WORD EMBEDDINGS IN ROMANIAN FOR THE RETAIL BANKING DOMAIN, In Proceedings of the IE 2020 International Conference</p>	BDI	2
	<p>Găman, M, Ionescu, RT. The unreasonable effectiveness of machine learning in Moldavian versus Romanian dialect identification. Int J Intell Syst. 2022; 37: 4928- 4966. doi:10.1002/int.22746</p>	ISI	4
	<p>Tufiş, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijccc.2022.1.4641</p>	ISI	4
	<p>Barbu, Eduard, and Verginica Barbu Mititelu. "Evaluating Computational Models of Similarity against a Human Rated Dataset." Baltic Journal of Modern Computing 10.3 (2022): 295-306.</p>	BDI	2
	<p>C. M. Popescu, C. Rusu and L. Grama, "Word Embeddings for Romanian Language and Their Use for Synonyms Detection," 2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), 2021, pp. 151-155, doi: 10.1109/SpeD53181.2021.9587432.</p>	BDI	2
<p>Păis, Vasile and Tufiş, Dan. More Romanian word embeddings from the</p>	<p>Tufiş, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1).</p>	ISI	4

<p>RETEROM project. In Proceedings of the International Conference on Linguistic Resources and Tools for Processing Romanian Language - CONSILR. pp. 91-100, 2018</p>	<p>doi:10.15837/ijcc.2022.1.4641</p>		
<p>Boroş, Tiberiu and Dumitrescu, Ştefan Daniel and Păiş, Vasile. Tools and resources for Romanian text-to-speech and speech-to-text applications. In Proceedings of the 15th International Conference on Human Computer Interaction (RoCHI 2018), pp 46-53, 2018</p>	<p>Verginica Barbu Mititelu, Dan Tufiş, Elena Irimia, The Reference Corpus of the Contemporary Romanian Language (CoRoLa), LREC 2018, L18-1189, 2018</p> <p>A. STAN, "Input Encoding for Sequence-to-Sequence Learning of Romanian Grapheme-to-Phoneme Conversion," 2019 International Conference on Speech Technology and Human-Computer Dialogue (SpED), 2019, pp. 1-6, doi: 10.1109/SPED.2019.8906639.</p> <p>Maria Nuţu, Deep Learning Approach for Automatic Romanian Lemmatization, Procedia Computer Science, Volume 192, 2021, Pages 49-58, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2021.08.006.</p> <p>Tufiş, D. (2022). Romanian Language Technology – a view from an academic perspective. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 17(1). doi:10.15837/ijcc.2022.1.4641</p> <p>Lörincz, B., Irimia, E., Stan, A., & Barbu Mititelu, V. (2023). RoLEX: The development of an extended Romanian lexical dataset and its evaluation at predicting concurrent lexical information. Natural Language Engineering, 29(3), 720-745. doi:10.1017/S1351324922000419</p> <p>Nicolae, C. D., Yadav, R. K., & Tufiş, D. (2023). Evaluation of Language Models on Romanian XQuAD and RoITD datasets. INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL, 18(1).</p> <p>Tăiatu, I. M., Avram, A. M., Cercel, D. C., & Pop, F. (2023, June). RoBERTweet: A BERT Language Model for Romanian Tweets. In International Conference on Applications of Natural Language to</p>	<p>ISI</p> <p>BDI</p> <p>ISI</p> <p>ISI</p> <p>ISI</p> <p>ISI</p>	<p>2.67</p> <p>1.33</p> <p>2.67</p> <p>2.67</p> <p>2.67</p> <p>2.67</p> <p>1.14</p> <p>0.57</p>
<p>Andrei-Marius Avram, Darius Catrina, Dumitru-Clementin Cercel, Mihai Dascalu, Traian Rebedea, Vasile Pais, and Dan Tufiş. 2022. Distilling the Knowledge of Romanian BERTs Using Multiple</p>		<p>BDI</p>	<p>0.57</p>

Teachers. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pages 374–384, Marseille, France. European Language Resources Association.	Information Systems (pp. 577-587). Cham: Springer Nature Switzerland.		
	Echim, S. V., Smădu, R. A., Avram, A. M., Cercel, D. C., & Pop, F. (2023, June). Adversarial Capsule Networks for Romanian Satire Detection and Sentiment Analysis. In International Conference on Applications of Natural Language to Information Systems (pp. 428-442). Cham: Springer Nature Switzerland.	BDI	0.57
	Andrei Avram, Verginica Barbu Mititelu, and Dumitru-Clementin Cercel. 2023. Romanian Multiword Expression Detection Using Multilingual Adversarial Training and Lateral Inhibition. In Proceedings of the 19th Workshop on Multiword Expressions (MWE 2023), pages 7–13, Dubrovnik, Croatia. Association for Computational Linguistics.	BDI	0.57

Total A3.1.1. Citări în cărți, reviste și volume ale unor manifestări științifice (cărți, ISI): 189

Total A3.1.2. Citări în cărți, reviste și volume ale unor manifestări științifice (BDI): 66.94

A3.3. Membru în comitetele de organizare ale manifestărilor științifice internaționale indexate BDI

Conferință workshop	Link	Punctaj
The 13th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language, 2018	https://profs.info.uaic.ro/~consilr/2018/index.php/comitetul-de-program/	6
RANLP 2021	https://ranlp.org/ranlp2021/pc.php	6
LivingNER @ IberLef 2022	https://temu.bsc.es/livingner/2022/01/28/scientific-committee/	6
The 15th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language, 2020	https://conferences.info.uaic.ro/consilr/previous-editions.html	6
The 16th Edition of the International Conference on	https://conferences.info.uaic.ro/consilr/previous-editions.html	6

Linguistic Resources and Tools for Processing Romanian Language, 2021		
The 17th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language, 2022	https://conferences.info.uaic.ro/consilr/previous-editions.html	6
The 18th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language, 2023	https://conferences.info.uaic.ro/consilr/previous-editions.html	6
The 19th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language, 2024	https://conferences.info.uaic.ro/consilr/committees.html	6
The 6th International Conference Computational Linguistics in Bulgaria, 2024	https://dcl.bas.bg/clib/committees/	6

Total A3.3. 54

A3.4. Premii in domeniul conferite de Academia Romana, ASTR, AOSR, sau premii internationale de prestigiu

Premiu	Punctaj
Premiul „Gh. Cartianu” al Academiei Române	15

Total A3.4. 15

Total A3. 324.94

Total A: 1198.73

Articole în reviste Q1,Q2 , conform ediției JCR-2022 (iunie 2023),
<https://uefiscdi.gov.ro/resource-866009-zone.iunie.2023.jif.pdf>

- (Q1) **Păis, Vasile** and Tufiș, Dan (2022). **Capitalization and punctuation restoration: a survey**. In: Artificial Intelligence Review 55.3, pp. 1681–1722. URL: <https://doi.org/10.1007/s10462-021-10051-x>.
- (Q1) **Păis, Vasile** and Ion, Radu and Irimia, Elena and Barbu Mîtitelu, Verginica and Badea, Valentin and Tufiș, Dan (2024). **System for the anonymization of Romanian jurisprudence**. Artificial Intelligence and Law. <https://link.springer.com/article/10.1007/s10506-024-09420-y>
- (Q1) Avram, Andrei-Marius and Mîtitelu, Verginica Barbu and **Păis, Vasile** and Cercei, Dumitru-Clementin and Trăușan-Matu, Ștefan (2023). **Multilingual Multiword Expression Identification Using Lateral Inhibition and Domain Adaptation**. Mathematics. <https://www.mdpi.com/2227-7390/11/11/2548>
- (Q2) **Păis, Vasile** and Mitrofan, Maria and Gasan, Carol Luca and Ianov, Alexandru and Ghiță, Corvin and Coneschi, Vlad Silviu and Onuț, Andrei (2024). **LegalNERo: A linked corpus for named entity recognition in the Romanian legal domain**. SemanticWeb Journal. <https://content.iopress.com/articles/semantic-web/sw233351>
- (Q2) **Păis, Vasile** and Barbu Mîtitelu, Verginica and Irimia, Elena and Ion, Radu and Tufiș, Dan (2024). **Underrepresented Speech Dataset from Open Data: Case Study on the Romanian Language**. Appl. Sci. 2024, 14(19), 9043; <https://doi.org/10.3390/app14199043>
- (Q2) **Păis, Vasile** and Ion, Radu and Avram, Andrei-Marius and Mitrofan, Maria and Tufiș, Dan (2021). **In-depth evaluation of Romanian natural language processing pipelines**. In Romanian Journal of Information Science and Technology (ROMJIST), vol. 24, no. 4, pp. 384–401, <https://www.romjist.ro/full-texts/paper700.pdf>.

Nr.Crt.	Domeniul de activitate	Punctaj
A1	Activitatea didactică / profesională (A1)	105.68
A2	Activitatea de cercetare (A2)	768.11
A3	Recunoașterea impactului activității (A3)	324.94
Total (A)		1198.73

Subcategori		Punctaj
A1.1.1-A1.1.2	Cărți de specialitate	2
A2.1	Articole in reviste cotate ISI si in volumele unor manifestări științifice indexate ISI proceedings	33, din care 6 în reviste cotate ISI Q1 sau Q2
A2.4.1	Granturi / proiecte de cercetare câștigate prin competiție (Director / Responsabil partener)	4
A3.1.1	Număr de citări in cărți, reviste cotate ISI si volume ale unor manifestări științifice ISI (WOS)	67
	Factor de impact ISI cumulat pentru publicații	26.8