

**Nume prenume: BODÓ Zalán-Péter**  
**Fișă de verificare a standardelor minime CNATDCU**

| <b>Criterii CNATDCU (cf. anexelor la OM 6129/20.12.2016)</b> |  |  |
|--|--|--|
|  | <b>Standarde minime</b>  | <b>Punctaj obținut</b>                     |
| <b>Producția științifică</b>                                 | <i>Conferențiar: 32 (<math>A^*+A+B \geq 16</math>)<br/>Profesor: 56 (<math>A^*+A \geq 24</math>, <math>A^*+A+B \geq 40</math>)</i> | 60.28<br>( $A^*+A+B = 44$ , $A^*+A = 24$ ) |
| <b>Impactul rezultatelor</b>                                 | <i>Conferențiar: 48 (<math>A^*+A+B \geq 12</math>)<br/>Profesor: 120 (<math>A^*+A+B \geq 40</math>)</i>                            | 400<br>( $A^*+A+B = 356$ )                 |
| <b>Performanța academică</b>                                 | <i>Conferențiar: 36<br/>Profesor: 60 (minim un proiect coordonat)</i>  | 63   |
|  |  | <b>Total: 523.28</b>                       |



|   |           |
|---|-----------|
| Zalán Bodó, Lehel Csató. Improving Kernel Locality-Sensitive Hashing Using Pre-Images and Bounds. In Proceedings of <b>IJCNN</b> , 2012, pp. 2710-2717. | 8         |
| <b>Punctaj total lucrări categoria A</b>  | <b>24</b> |

| Lucrări categoria B  |         |
|--|---------|
| Autori, Titlul publicatiei, Referința Bibliografică  | Punctaj |
| Attila Mester, Zalán Bodó. Validating static call graph-based malware signatures using community detection methods. <b>ESANN</b> , 2021, pp. 429-434.  | 4       |
| Anna Kiss, Csaba Sulyok, Zalán Bodó. Region Prediction from Hungarian Folk Music Using Convolutional Neural Networks. Artificial Neural Networks and Machine Learning – <b>ICANN</b> 2019: Text and Time Series, Lecture Notes in Computer Science, vol. 11730, Springer, Cham, 2019, pp. 581-594.   | 4       |
| Zalán Bodó, Bipin Indurkha. Software categorization using low-level distributional features. New Trends in Intelligent Software Methodologies, Tools and Techniques. (Proceedings of the 16th International Conference on Intelligent Software Methodologies, Tools, and Techniques ( <b>SOMET</b> ), September 26--28, Kitakyushu, Japan.) Frontiers in Artificial Intelligence and Applications, vol. 297, IOS Press, 2017, pp. 88-98. | 4       |
| Zalán Bodó, Lehel Csató. Augmented hashing for semi-supervised scenarios. In Proceedings of the 22th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning ( <b>ESANN</b> ), 2014, pp. 53-58.  | 4       |
| Zalán Bodó, Lehel Csató. Linear Spectral Hashing. In Proceedings of the 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning ( <b>ESANN</b> ), 2013, pp. 303-308.  | 4       |

|  |           |
|--|-----------|
|  |           |
| <b>Punctaj total lucrări categoria B</b> | <b>20</b> |

| <b>Lucrări categoria C</b>  |          |
|---|----------|
| Autori, Titlul publicatiei, Referința Bibliografică   | Punctaj  |
| Adél Bajcsi, Anna Bajcsi, Szabolcs Pével, Ábel Portik, Csanád Sándor, Annamária Szenkovits, Orsolya Vas, Zalán Bodó, Lehel Csató. Comparative Study of Interpretable Image Classification Models. <b>Infocommunications Journal</b> , Special Issue on Applied Informatics, 2023, pp. 20-26.                              | 2/7=0.28 |
| Attila Mester, Zalán Bodó. Malware classification based on graph convolutional neural networks and static call graph features. <b>IEA/AIE 2022: Advances and Trends in Artificial Intelligence. Theory and Practices in Artificial Intelligence</b> , pp. 528-539, Springer, Cham.  | 2        |
| Zalán Bodó. Fake News Detection Without External Knowledge. In Proceedings of the International Conference on Modelling and Development of Intelligent Systems (MDIS) 2020, Communications in Computer and Information Science series (CCIS), volume 1341, pages 202-221, <b>Springer</b> International Publishing, 2021. | 2        |
| Zalán Bodó. A CiteSeerX-based dataset for record linkage and metadata extraction. Proceedings of the 2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing ( <b>SYNASC 2018</b> , Timisoara, Romania, 20-23 September 2018), pp. 230-236, IEEE, 2018.                             | 2        |
| Zalán Bodó, Zsolt Minier, Lehel Csató. Active Learning with Clustering. JMLR Workshop and Conference Proceedings: Volume 16, (Active Learning and Experimental Design workshop, May 16, 2010, Sardinia, Italy) 2011, pp. 127-139. 1 pct. <b>AISTATS</b> (B cat. Conf. - ERA 2010) <b>Workshop</b> paper.                  | 4/2=2    |

|  |              |
|--|--------------|
| Zalán Bodó, Lehel Csató. Hierarchical and Reweighting Cluster Kernels for Semi-Supervised Learning. <b>Int. J. of Computers, Communications &amp; Control</b> , Vol. V (2010), No. 4, pp. 469-476.   | 2            |
| Zalán Bodó. Hierarchical cluster kernels for supervised and semi-supervised learning. In Proceedings of the IEEE 4th International Conference on Intelligent Computer Communication and Processing ( <b>ICCP</b> ), 2008, pp. 9-16.                                | 2            |
| Zsolt Minier, Zalán Bodó, Lehel Csató. Wikipedia-based kernels for text categorization. In Proceedings of the 9th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing ( <b>SYNASC'07</b> ), Timisoara, Romania, 2007, pp. 157-164. | 2            |
| Zsolt Minier, Zalán Bodó, Lehel Csató. Segmentation-based feature selection for text categorization. In Proceedings of the IEEE 2nd International Conference on Intelligent Computer Communication and Processing ( <b>ICCP</b> ), pp. 53-59, September 2006.      | 2            |
| <b>Punctaj total lucrări categoria C</b>   | <b>16.28</b> |

|   |            |
|---|------------|
| <b>Nume, Prenume: Bodó Zalán-Péter</b><br>Impactul rezultatelor |            |
| <b>Punctaj total citări</b>                                     | <b>400</b> |
| <b>Punctaj citări din forumuri de tip A*+A+B</b>                | <b>356</b> |

| Lucrare citată  | Citări  | Categoria CNATDCU | Punctaj |
|---|---|-------------------|---------|
| László Antal, Zalán Bodó. Feature axes orthogonalization in semantic face editing. In proceedings of ICCP 2021 (IEEE 17th International Conference on Intelligent Computer Communication and Processing) October 28-30, 2021, Cluj-Napoca, Romania (online event), pp. 163-169. | Zhou, Chenliang, Fangcheng Zhong, and Cengiz Öztireli. "CLIP-PAE: Projection-Augmentation Embedding to Extract Relevant Features for a Disentangled, Interpretable and Controllable Text-Guided Face Manipulation." <i>ACM SIGGRAPH 2023 Conference Proceedings</i> . 2023. | A*                | 12      |

| Lucrare citată  | Citări   | Categoria CNATDCU | Punctaj |
|---|--|-------------------|---------|
| Attila Mester, Zalán Bodó. Validating static call graph-based malware signatures using community detection methods. In proceedings of ESANN 2021 (29th European | Amira, Abdelouahab, et al. "A survey of malware analysis using community detection algorithms." <i>ACM Computing Surveys</i> (2023). | A*                | 12      |

|   |  |
|---|--|
| Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning), 6-8 October 2021, Bruges, Belgium (online event), pp. 429-434. |  |
|---|--|

| Lucrare citată  | Citări  | Categoria CNATDCU | Punctaj |
|---|---|-------------------|---------|
| Csaba Sulyok, Christopher Harte, Zalán Bodó. On the Impact of Domain-specific Knowledge in Evolutionary Music Composition. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO), 2019, pp. 188-197. | López-Montes, José, Miguel Molina-Solana, and Waldo Fajardo. "GenoMus: Representing Procedural Musical Structures with an Encoded Functional Grammar Optimized for Metaprogramming and Machine Learning." <i>Applied Sciences</i> 12.16 (2022): 8322. | C                 | 2       |
|   | Wen, Yu-Wei, and Chuan-Kang Ting. "Recent Advances of Computational Intelligence Techniques for Composing Music." <i>IEEE Transactions on Emerging Topics in Computational Intelligence</i> (2022).   | B                 | 4       |

| Lucrare citată  | Citări  | Categoria CNATDCU | Punctaj |
|---|---|-------------------|---------|
| Zalán Bodó, Bipin Indurkha. Software categorization using low-level distributional features. <i>New Trends in Intelligent Software Methodologies, Tools and Techniques</i> . (Proceedings of the 16th International Conference on Intelligent Software Methodologies, | Auch, Maximilian, et al. "Similarity-based analyses on software applications: A systematic literature review." <i>Journal of Systems and Software</i> 168 (2020): 110669. | B                 | 4       |

|  |  |
|--|--|
| Tools, and Techniques, September 26--28, Kitakyushu, Japan.)<br>Frontiers in Artificial Intelligence and Applications, vol. 297, IOS Press, 2017, pp. 88-98. |  |
|--|--|

| Lucrare citată  | Citări   | Categoria CNATDCU | Punctaj |
|---|--|-------------------|---------|
| Zalán Bodó, Lehel Csató. A Hybrid Approach for Scholarly Information Extraction. <i>Studia Universitatis Babes-Bolyai Informatica</i> , Vol. 62, No. 2, 2017, pp. 5-16. | Zaman, Gohar, et al. "An ontological framework for information extraction from diverse scientific sources." <i>IEEE Access</i> 9 (2021): 42111-42124.                    | B                 | 4       |
|   | Zaman, Gohar, et al. "Information extraction from semi and unstructured data sources: A systematic literature review." <i>ICIC Express Letters</i> 14.6 (2020): 593-603. | C                 | 2       |

| Lucrare citată  | Citări  | Categoria CNATDCU | Punctaj |
|---|---|-------------------|---------|
| Zalán Bodó, Lehel Csató. Linear Spectral Hashing. <i>Neurocomputing</i> , Volume 141, 2 October 2014, pp. 117-123.<br><br>& | Di Wang, Xinbo Gao, Xiumei Wang. Semi-supervised constraints preserving hashing. <i>Neurocomputing</i> , Volume 167, 1 November 2015, Pages 230–242.          | A                 | 8       |
|   | Tao Yao, Xiangwei Kong, Haiyan Fu, Qi Tian. Semantic consistency hashing for cross-modal retrieval. <i>Neurocomputing</i> , Available online 20 February 2016 | B                 | 4       |

|   |  |    |    |
|---|--|----|----|
| <p>Zalán Bodó, Lehel Csató. Linear Spectral Hashing. In Proceedings of the 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2013, pp. 303-308.</p> | <p>Ribeiro, Bernardete, and Ning Chen. "Financial credit risk assessment via learning-based hashing." <i>Intelligent Decision Technologies</i> 11.2 (2017): 177-186.</p>   | C  | 2  |
|   | <p>Hongjian Zha, Shixiong Xia, Rui Yao, Qiang Niu and, Yong Zhou. Real-time video fusion based on multistage hashing and hybrid transformation with depth adjustment, <i>J. Electron. Imaging.</i> 24(6), 063023 (Dec 30, 2015). <a href="http://dx.doi.org/10.1117/1.JEI.24.6.063023">http://dx.doi.org/10.1117/1.JEI.24.6.063023</a></p> | C  | 2  |
|   | <p>Bernardete Ribeiro, Ning Chen. Hashing for Financial Credit Risk Analysis. <i>Neural Information Processing</i>, Volume 8835 of the series <i>Lecture Notes in Computer Science</i>, pp. 395-403 (21st International Conference, ICONIP 2014, Kuching, Malaysia, November 3-6, 2014. Proceedings, Part II).</p>                         | A  | 8  |
|   | <p>Karbil, Loubna, Imane Daoudi, and Hicham Medromi. A Comparative Experimental Study of Spectral Hashing. <i>Advances in Ubiquitous Networking 2</i>. Springer Singapore, 2017. 445-456.</p>  | C  | 2  |
|   | <p>Yao, Tao, et al. "Supervised Coarse-to-Fine Semantic Hashing for cross-media retrieval." <i>Digital Signal Processing</i> 63 (2017): 135-144.</p>   | B  | 4  |
|   | <p>Chi, Lianhua, and Xingquan Zhu. "Hashing Techniques: A Survey and Taxonomy." <i>ACM Computing Surveys (CSUR)</i> 50.1 (2017): 11.</p>   | A* | 12 |
|   | <p>Ding, Guiguang, et al. "Large-scale image retrieval with Sparse Embedded Hashing." <i>Neurocomputing</i> (2017).</p>  | A  | 8  |

| Lucrare citată   | Citări  | Categoria CNATDCU | Punctaj |
|--|---|-------------------|---------|
| <p>Zalán Bodó, Zsolt Minier, Lehel Csató. Active Learning with Clustering. Active Learning Challenge: Challenges in Machine Learning, Volume 6, Microtome Publishing, 2012, pp. 141-154.</p> <p>&amp;</p> <p>Zalán Bodó, Zsolt Minier, Lehel Csató. Active Learning with Clustering. JMLR Workshop and Conference Proceedings: Volume 16, (Active Learning and Experimental Design workshop, May 16, 2010, Sardinia, Italy) 2011, pp. 127-139.</p> | <p>Maheshwary, Saket, and Saurabh Sohoney. "Learning geolocation by accurately matching customer addresses via graph based active learning. WWW 2023.</p>   | A*                | 12      |
|  | <p>Hassan, Sabit, and Malihe Alikhani. "D-CALM: A Dynamic Clustering-based Active Learning Approach for Mitigating Bias." arXiv preprint arXiv:2305.17013 (2023). (Findings of the ACL)</p>   | A*                | 12      |
|  | <p>Margatina, Katerina, and Nikolaos Aletras. "On the Limitations of Simulating Active Learning." <i>arXiv preprint arXiv:2305.13342</i> (2023).<br/>(Published in Findings of the Association for Computational Linguistics: ACL 2023)</p> | A*                | 12      |
|  | <p>Wanyan, Yuyang, et al. "Active Exploration of Multimodal Complementarity for Few-Shot Action Recognition." <i>Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition</i>. 2023.</p>                           | A*                | 12      |
|  | <p>Rana, Aayush J., and Yogesh S. Rawat. "Hybrid Active Learning via Deep Clustering for Video Action Detection." <i>Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition</i>. 2023.</p>                       | A*                | 12      |
|  | <p>Treviso, Marcos, et al. "Efficient methods for natural language processing: A survey." <i>Transactions of the Association for Computational Linguistics</i> 11 (2023):</p>   | C                 | 2       |

|  |  |    |    |
|--|--|----|----|
|  | 826-860.   |    |    |
|  | Kwak, Beong-woo, et al. "TrustAL: Trustworthy Active Learning using Knowledge Distillation." arXiv preprint arXiv:2201.11661 (2022). AAAI 2022.  | A* | 12 |
|  | Katerina Margatina, Loic Barrault, and Nikolaos Aletras. 2022. On the Importance of Effectively Adapting Pretrained Language Models for Active Learning. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 825–836, Dublin, Ireland. Association for Computational Linguistics. | A* | 12 |
|  | Xinyu Hua and Lu Wang. 2022. Efficient Argument Structure Extraction with Transfer Learning and Active Learning. In Findings of the Association for Computational Linguistics: ACL 2022, pages 423–437, Dublin, Ireland. Association for Computational Linguistics.  | A* | 12 |
|  | Vaaras, Einari, Manu Airaksinen, and Okko Räsänen. "Analysis of Self-Supervised Learning and Dimensionality Reduction Methods in Clustering-Based Active Learning for Speech Emotion Recognition." arXiv preprint arXiv:2206.10188 (2022). Interspeech 2022.   | A  | 8  |
|  | Zhang, Shujian, et al. "ALLSH: Active Learning Guided by Local Sensitivity and Hardness." arXiv preprint arXiv:2205.04980 (2022). NAACL 2022.  | A  | 8  |
|  | Katerina Margatina, Giorgos Vernikos, Loïc Barrault, and Nikolaos Aletras. 2021. Active Learning by Acquiring Contrastive Examples. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pages 650–  | A  | 8  |

|  |   |             |   |
|--|---|-------------|---|
|  | 663, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.  |             |   |
|  | Prabhu, Sumanth, Moosa Mohamed, and Hemant Misra. "Multi-class text classification using BERT-based active learning." arXiv preprint arXiv:2104.14289 (2021). DaSH@KDD'21   | A* workshop | 8 |
|  | Sreenivasaiyah, Deepthi, Johannes Otterbach, and Thomas Wollmann. "MEAL: Manifold Embedding-based Active Learning." 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). IEEE Computer Society, 2021. ICCV workshop | A* workshop | 8 |
|  | A comprehensive study of class incremental learning algorithms for visual tasks<br>E Belouadah, A Popescu, I Kanellos - Neural Networks, 2020 - Elsevier  | A           | 8 |
|  | Nuclear discrepancy for single-shot batch active learning<br>TJ Viering, JH Krijthe, M Loog - Machine Learning, 2019 - Springer   | A           | 8 |
|  | Batch Active Learning With Two-Stage Sampling. R Luo, X Wang - IEEE Access, 2020  | A           | 8 |
|  | Batch mode active learning for node classification in assortative and disassortative networks<br>S Ping, D Liu, B Yang, Y Zhu, H Chen, Z Wang - IEEE Access, 2017   | A           | 8 |

|  |   |   |   |
|--|---|---|---|
|  | Fabian L. Wauthier, Nebojsa Jojic, Michael I. Jordan. Active Spectral Clustering via Iterative Uncertainty Reduction. SIGKDD'12, 2012, pp. 1339-1347.   | A | 8 |
|  | Ngomo, Axel-Cyrille Ngonga, Klaus Lyko, and Victor Christen. COALA–Correlation-Aware Active Learning of Link Specifications. The Semantic Web: Semantics and Big Data. Springer Berlin Heidelberg, 2013. 442-456. | C | 2 |
|  | Shayok Chakraborty, Vineeth Balasubramanian, Sethuraman Panchanathan. Batch Mode Active Learning for Multimedia Pattern Recognition. IEEE International Symposium on Multimedia (ISM), 2012, pp. 489-490.         | C | 2 |
|  | Fu, Chunjiang, Yang, Yupu. Low density separation as a stopping criterion for active learning SVM. Intelligent Data Analysis, vol. 19, no. 4, pp. 727-741, 2015   | C | 2 |
|  | Fu, Chun-Jiang, Yang, Yu-Pu. A batch-mode active learning SVM method based on semi-supervised clustering. Journal: Intelligent Data Analysis, vol. 19, no. 2, pp. 345-358, 2015.                                  | C | 2 |
|  | Vu, Viet-Vu, and Nicolas Labroche. "Active seed selection for constrained clustering." Intelligent Data Analysis 21.3 (2017): 537-552.  | C | 2 |

| Lucrare citată | Citări | Categoria CNATDCU | Punctaj |
|----------------|--------|-------------------|---------|
|----------------|--------|-------------------|---------|

|   |   |   |   |
|---|---|---|---|
| Zalán Bodó, Lehel Csató.<br>Hierarchical and Reweighting<br>Cluster Kernels for Semi-Supervised<br>Learning. Int. J. of Computers,<br>Communications & Control, Vol. V<br>(2010), No. 4, pp. 469-476. | Ma, Jun, et al. "Safe semi-supervised learning for pattern classification."<br>Engineering Applications of Artificial Intelligence 121 (2023): 106021.  | A | 8 |
|   | Madeleine Seeland, Andreas Karwath, Stefan Kramer. A structural cluster kernel<br>for learning on graphs. KDD '12 Proceedings of the 18th ACM SIGKDD<br>International Conference on Knowledge Discovery and Data Mining, pp. 516–<br>524. | A | 8 |

| Lucrare citată   | Citări  | Categoria<br>CNATDCU | Punctaj |
|--|---|----------------------|---------|
| Zalán Bodó. Hierarchical cluster<br>kernels for supervised and semi-<br>supervised learning. In Proceedings<br>of the IEEE 4th International<br>Conference on Intelligent Computer<br>Communication and Processing,<br>2008, pp. 9-16. | Madeleine Seeland, Andreas Karwath, Stefan Kramer. A structural cluster kernel<br>for learning on graphs. KDD '12 Proceedings of the 18th ACM SIGKDD<br>International Conference on Knowledge Discovery and Data Mining, pp. 516–<br>524. | A                    | 8       |

| Lucrare citată  | Citări   | Categoria<br>CNATDCU | Punctaj |
|---|--|----------------------|---------|
| Zalán Bodó, Zsolt Minier. On<br>supervised and semi-supervised k- | Zhang, Jun, et al. "Enhanced semi-supervised ensemble machine learning<br>approach for earthwork construction simulation activity sequence automatically | C                    | 2       |

|   |  |          |          |
|---|--|----------|----------|
| <p>nearest neighbor algorithms. In Proceedings of the 7th Joint Conference on Mathematics and Computer Science, Cluj-Napoca, 2008. (Studia Universitatis Babes-Bolyai, Series Informatica, Volume LIII, Number 2)</p> | <p>updating driven by weather data." <i>Geological Journal</i> 58.6 (2023): 2231-2253.</p>   |          |          |
|   | <p>Ajita Satheesh, Ravindra Patel. Dynamic Nearest Neighbours Classifier For Integrated Data Using Object Oriented Concept Generalization. <i>IJSSST</i>, Vol. 11, No. 1, 2010, pp. 35–40.</p> | <p>C</p> | <p>2</p> |

| Lucrare citată  | Citări  | Categoria CNATDCU | Punctaj  |
|---|---|-------------------|----------|
| <p>Zalán Bodó, Zsolt Minier, Lehel Csató. Text categorization experiments using Wikipedia. In Proceedings of the 1st 'Knowledge Engineering: Principles and Techniques' Conference, pp. 66-72, Cluj-Napoca, Romania, 2007. (Studia Universitatis Babes-Bolyai, Series Informatica, Volume LII, Special Issue KEPT-2007)</p> | <p>COURSEY, Kino; MIHALCEA, Rada; MOEN, William. Using encyclopedic knowledge for automatic topic identification. In: Proceedings of the Thirteenth Conference on Computational Natural Language Learning. Association for Computational Linguistics, 2009. p. 210-218.</p> | <p>A</p>          | <p>8</p> |

| Lucrare citată   | Citări  | Categoria CNATDCU | Punctaj |
|--|---|-------------------|---------|
| <p>Zsolt Minier, Zalán Bodó, Lehel Csató. Wikipedia-based kernels for text categorization. In Proceedings of the 9th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC'07), Timisoara, Romania, 2007, pp. 157-164.</p> | <p>Anna Huang, David Milne, Eibe Frank, and Ian H. Witten. Clustering documents with active learning using Wikipedia. In <i>Proc. of the 8th IEEE International Conference on Data Mining</i>, Pisa, Italy, pages 839-844. IEEE Computer Society, 2008.</p> | A*                | 12      |
|  | <p>Anna Huang, David Milne, Eibe Frank, and Ian H. Witten. Clustering Documents using a Wikipedia-based Concept Representation, <i>Proc. of the Pacific-Asia Conference on Knowledge Discovery and Data Mining</i>, Bangkok, Thailand, April 2009.</p>      | A                 | 8       |
|  | <p>Zeno Gantner, Lars Schmidt-Thieme. Automatic Content-based Categorization of Wikipedia Articles, Proceedings of the 2009 Workshop on the People's Web Meets NLP, ACL-IJCNLP 2009, pages 32–37, Suntec, Singapore, 7 August 2009.</p>                     | A workshop        | 4       |
|  | <p>Olena Medelyan, David Milne, Catherine Legg, Ian H. Witten. Mining meaning from Wikipedia. <i>International Journal of Human-Computer Studies</i>, Volume 67, Issue 9 (September 2009), Pages 716-754, 2009. (<b>după CORE</b>)</p>                      | A                 | 8       |
|  | <p>Tamara Polajnar, Mark Girolami. Semi-supervised Prediction of Protein Interaction Sentences Exploiting Semantically Encoded Metrics. <i>Pattern Recognition in Bioinformatics</i>, Springer Berlin / Heidelberg, pages 270-281, 2009.</p>                | C                 | 2       |
|  | <p>Tamara Polajnar, Mark Girolami. Application of Lexical Topic Models to Protein</p>   | A* workshop       | 8       |

|  |   |   |   |
|--|---|---|---|
|  | Interaction Sentence Prediction, In NIPS 2009 Workshop on Applications for Topic Models: Text and Beyond, 2009.   |   |   |
|  | Bing Bai, Jason Weston, David Grangier, Ronan Collobert, Kunihiko Sadamasa, Yanjun Qi, Olivier Chapelle, Kilian Weinberger. Learning to Rank with (a Lot of) Word Features. Information Retrieval, Springer Netherlands, 2009. ( <b>după CORE</b> )                                 | B | 4 |
|  | Catarina Silva, Bernadete Ribeiro. Inductive Inference for Large Scale Text Classification. Kernel Approaches and Techniques. Series: Studies in Computational Intelligence, Vol. 255. Springer Berlin / Heidelberg, 2010.  | C | 2 |
|  | B. Bai, J. Weston, D. Grangier, R. Collobert, O. Chapelle, K. Weinberger. Supervised Semantic Indexing. The 18th ACM Conference on Information and Knowledge Management (CIKM), 2009.   | A | 8 |
|  | Milad Alemzadeh, Fakhri Karray. An Efficient Method for Tagging a Query with Category Labels using Wikipedia towards Enhancing Search Engine Results. Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2010, pp. 192–195. | C | 2 |
|  | Michael Hart, Pratyusa Manadhata, Rob Johnson. Text Classification for Data Loss Prevention. Lecture Notes in Computer Science, 2011, Volume 6794/2011, pp. 18–37.  | C | 2 |
|  | Sami Virpioja, Mari-Sanna Paukkeri, Abhishek Tripathi, Tiina Lindh-Knuutila, and Krista Lagus. Evaluating vector space models with canonical correlation analysis. Natural Language Engineering, Cambridge University Press, Vol. 18, Issue 03,                                     | A | 8 |

|  |   |   |   |
|--|---|---|---|
|  | 2012, pp. 399–436. (după CORE)  |   |   |
|  | Tamara Polajnar, Nitish Aggarwal, Kartik Asooja, Paul Buitelaar. Improving ESA with Document Similarity. Lecture Notes in Computer Science Volume 7814, 2013, pp. 582-593.  | C | 2 |
|  | Christoph Schaefer, Daniel Hienert, Thomas Gottron. Normalized Relevance Distance – A Stable Metric for Computing Semantic Relatedness over Reference Corpora. pp. 789-794, DOI 10.3233/978-1-61499-419-0-789, Series Frontiers in Artificial Intelligence and Applications, Volume 263: ECAI 2014. | A | 8 |
|  | Aibek Musaev, De Wang, Saajan Shridhar, Calton Pu. Fast Text Classification Using Randomized Explicit Semantic Analysis. Information Reuse and Integration (IRI), 2015 IEEE International Conference on.  | C | 2 |

| Lucrare citată   | Citări  | Categoria CNATDCU | Punctaj |
|--|---|-------------------|---------|
| Zalán-Péter Bodó. Maximal processor utilization in parallel quadtree-based fractal image compression on MIMD architectures, Studia Univ. Babeş-Bolyai, Informatica, Volume XLIX, Number 2, 2004. | Asati, Ranjita, M. M. Raghuwanshi, and Kavita R. Singh. "Fractal Image Coding-Based Image Compression Using Multithreaded Parallelization." Information and Communication Technology for Competitive Strategies (ICTCS 2021). Springer, Singapore, 2023. 559-569. | C                 | 2       |
|  | Nileshsingh V. Thakur, O. G. Kakde. Color Image Compression with Modified Fractal Coding on Spiral Architecture, Journal of Multimedia, Vol. 2, No. 4,  | C                 | 2       |

|  |  |
|--|--|
|  | August 2007.   |
|  | Meiqing Wang, Choi-Hong Lai. Grey video compression methods using fractals. C 2<br>International Journal of Computer Mathematics. Volume 84, Issue 11, 2007. |

**Nume Prenume: Bodó Zalán-Péter**  
**Performanța academică**

**Punctaj total** **63**

**Performanța academică**

Categorie

Punctaj

**Cărți și capitole de cărți: 4**

Bodó Zalán. Fordítóprogramok szerkesztése Flex és Bison segítségével. Societatea Muzeului Ardelean (Erdélyi Múzeum Egyesület), Cluj-Napoca, 2014. (carte, 100 pagini, editură nelistată)

D

2

Csató Lehel, Bodó Zalán. Neurális hálók és a gépi tanulás módszerei. Presa Universitară Clujeană, Cluj-Napoca, 2008. (carte, 179 pagini, editură nelistată)

D

2

**Editor Proceedings**

**Publicare curs universitar: 2**

|  |                              |   |
|--|------------------------------|---|
| Curs universitar <i>Formális nyelvek és fordítóprogramok</i> (Limbae formale și tehnici de compilare - în lb. maghiară), care conține: un curs scris (100 de pagini), slide-uri pentru curs (> 200 de pagini), probleme de laborator (15 pagini), exerciții pentru seminar (13 pagini):<br><a href="https://www.cs.ubbcluj.ro/~zbodo/?a=courses/fnyf/fnyf">https://www.cs.ubbcluj.ro/~zbodo/?a=courses/fnyf/fnyf</a> |                              | 2 |
| <b>Director/Editor revista</b>   |                              |   |
| <b>Director/Coordonator/Responsabil grant: 20</b>  |                              |   |
| Director proiect de cercetare finanțat de MTA (Academia Maghiară de Științe) Domus, Magyar könnyűzenei dalszövegek elemzése statisztikai természetesnyelv-feldolgozási módszerekkel (Analyzing Hungarian popular music lyrics using statistical natural language processing methods), grant ID 91/22/2023/HTMT   | 400000 Ft = 1071<br>EUR      | 2 |
| Director proiect de cercetare finanțat de MTA (Academia Maghiară de Științe) Domus, Adatbányászati módszerek a vírusselemezésben (Data mining methods in malware analysis), grant ID 86/18/2022/HTMT   | 750000 forinti = 2000<br>EUR | 2 |
| Membri proiect PCCA "Metode de îmbunătățire a evaluării cercetării prin analiza rețelelor științifice", 2012-2016, cod: PN-II-PT-PCCA- 2011-3.2-0895   | 2999500 lei = 674000<br>EUR  | 5 |
| Membri proiect CNCSIS-TE "Metode neparametrice în instruirea automată a mașinilor: aplicații în robotică și analiza datelor", 2011-2014, cod: PN-II-RU-TE-2011-3-0278  | 749400 lei = 168400<br>EUR   | 3 |
| Membri proiect CNMP "Metode de control al roboților autonomi folosind rețele neuronale cu pulsuri", 2007-2011, cod: NEUROBOT 11039/10.04.2007  | 172000 lei = 38650<br>EUR    | 1 |

|  |  |     |
|--|--|-----|
| Coordonator proiect Tineri Doctoranzi "Mașini de învățare în categorizarea documentelor", 2007-2008, cod: TD-35  | 25140 lei = 5650 EUR                               | 2   |
| Membru proiect de cercetare MOLINGV NKFP-2/0024/2005 (finanțat de Agenția Națională de Cercetare și Dezvoltare din Ungaria, prin Programul Național de Cercetare și Dezvoltare; institute participante: MTA SZTAKI (MLHCI group and Data Mining and Web Search Group), ELTE (Department of Computer Science, Physics of Complex Systems, BME (Math. Institute Department of Stochastic Analysis), Research Institute for Linguistics (HAS), MTA SZFKI, Omega Consulting, Pont Rendszerház), <i>Language Miner</i> , Omega Consulting Ltd., Ungaria ( <a href="https://sztaki.hun-ren.hu/en/projects/nyelvbanyasz">https://sztaki.hun-ren.hu/en/projects/nyelvbanyasz</a> ) (2005-2006, 5 luni) | aprox. 100 000 000 Ft<br>* 0.004 =<br>= 400000 EUR | 4   |
| Membru proiect de cercetare Fraktálfüggvények és alkalmazásaik (Funcții fractale și aplicațiile lor), Universitatea Sapientia, Cluj, coordonator proiect: dr. Anna Soós. (2003-2004)   | < 50000 EUR  | 1   |
| <b>Membru în Comitet științific: 9</b>   |  |     |
| The 36th International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems (IEA/AIE 2023) ( <a href="http://www.ieaaie2023.com/committee.html">http://www.ieaaie2023.com/committee.html</a> )  | C  | 1   |
| The 35th International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems (IEA/AIE 2022) ( <a href="https://ieaaie2022.wordpress.com/committee/">https://ieaaie2022.wordpress.com/committee/</a> )  | C  | 1   |
| FROM (The 4th Working Formal Methods Symposium) 2020 ( <a href="http://www.cs.ubbcluj.ro/from2020/committees/">http://www.cs.ubbcluj.ro/from2020/committees/</a> )   | D  | 0.5 |
| ICAI (The 11th International Conference on Applied Informatics) 2020 ( <a href="https://icai.uni-eszterhazy.hu/2020/">https://icai.uni-eszterhazy.hu/2020/</a> )   | D  | 0.5 |
| PCIT (PANNONIAN CONFERENCE ON ADVANCES IN INFORMATION TECHNOLOGY) 2020   | D  | 0.5 |

|  |          |     |
|--|----------|-----|
| <a href="https://pcit2020.mik.uni-pannon.hu/index.php/hu/tudomanyos-bizottsag.html">https://pcit2020.mik.uni-pannon.hu/index.php/hu/tudomanyos-bizottsag.html</a>  |          |     |
| PCIT (PANNONIAN CONFERENCE ON ADVANCES IN INFORMATION TECHNOLOGY) 2019<br>( <a href="https://pcit2019.mik.uni-pannon.hu/index.php/hu/tudomanyos-bizottsag.html">https://pcit2019.mik.uni-pannon.hu/index.php/hu/tudomanyos-bizottsag.html</a> )  | D        | 0.5 |
| FMTÜ 2019 (XXIV. Fiatal Műszakiak Tudományos Ülésszaka, 2019 - INTERNATIONAL SCIENTIFIC CONFERENCE OF YOUNG ENGINEERS) ( <a href="https://eda.eme.ro/handle/10598/31247">https://eda.eme.ro/handle/10598/31247</a> )   | D        | 0.5 |
| FMTÜ 2015 (XX. Fiatal Műszakiak Tudományos Ülésszaka, 2015 - INTERNATIONAL SCIENTIFIC CONFERENCE OF YOUNG ENGINEERS) ( <a href="https://eda.eme.ro/handle/10598/28675">https://eda.eme.ro/handle/10598/28675</a> )   | D        | 0.5 |
| <i>Membru în comitetul științific al workshopului ECML/PKDD "ALRA: ACTIVE LEARNING IN REAL-WORLD APPLICATIONS", 2012</i> ( <a href="http://fr.nomao.com/labs/alra">http://fr.nomao.com/labs/alra</a> )   | A (ECML) | 4   |
| <b>Director/Membru Comitet de Organizare eveniment științific: 16</b>  |          |     |
| Membru comitet organizare + comitet științific: A Magyar Tudomány Napja Erdélyben, 14. Matematika és Informatika Alkalmazásokkal Konferencia, 2023. november 16-18.<br>( <a href="https://www.cs.ubbcluj.ro/~darvay/eme/mtne2023/">https://www.cs.ubbcluj.ro/~darvay/eme/mtne2023/</a> ) | D        | 1   |
| Membru comitet organizare + comitet științific: A Magyar Tudomány Napja Erdélyben, 13. Matematika és Informatika Alkalmazásokkal Konferencia, 2022. november 25-27.<br>( <a href="https://www.cs.ubbcluj.ro/~darvay/eme/mtne2022/">https://www.cs.ubbcluj.ro/~darvay/eme/mtne2022/</a> ) | D        | 1   |
| Organizator principal MaCS 2022 (24-27 November, 2022) ( <a href="https://www.cs.ubbcluj.ro/~macs/2022/">https://www.cs.ubbcluj.ro/~macs/2022/</a> )   | D        | 2   |

|   |   |   |
|---|---|---|
| <p>Membru comitet organizare + comitet științific: A Magyar Tudomány Napja Erdélyben, 12. Matematika és Informatika Alkalmazásokkal Konferencia, 2021. november 12-14.<br/> <a href="https://www.cs.ubbcluj.ro/~darvay/eme/mtne2021/">(https://www.cs.ubbcluj.ro/~darvay/eme/mtne2021/)</a></p> | D | 1 |
| <p>Membru comitet organizare + comitet științific: A Magyar Tudomány Napja Erdélyben, 11. Matematika és Informatika Alkalmazásokkal Konferencia, 2020. november 13-15.<br/> <a href="https://www.cs.ubbcluj.ro/~darvay/eme/mtne2020/">(https://www.cs.ubbcluj.ro/~darvay/eme/mtne2020/)</a></p> | D | 1 |
| <p>Membru comitet organizare + comitet științific: A Magyar Tudomány Napja Erdélyben, 10. Matematika és Informatika Alkalmazásokkal Konferencia, 2019. november 15-17.<br/> <a href="https://www.cs.ubbcluj.ro/~darvay/eme/mtne2019/">(https://www.cs.ubbcluj.ro/~darvay/eme/mtne2019/)</a></p> | D | 1 |
| <p>Membru în comitetul de organizare a conferinței MACS 2018 (<a href="http://www.cs.ubbcluj.ro/~macs/2018/">http://www.cs.ubbcluj.ro/~macs/2018/</a>)</p>  | D | 1 |
| <p>Membru comitet organizare + comitet științific: A Magyar Tudomány Napja Erdélyben, 9. Matematika és Informatika Alkalmazásokkal Konferencia, 2018. november 16-18.<br/> <a href="https://www.cs.ubbcluj.ro/~darvay/eme/mtne2018/">(https://www.cs.ubbcluj.ro/~darvay/eme/mtne2018/)</a></p>  | D | 1 |
| <p>Membru comitet organizare + comitet științific: A Magyar Tudomány Napja Erdélyben, 8. Matematika és Informatika Alkalmazásokkal Konferencia, 2017. november 3-5.<br/> <a href="https://www.cs.ubbcluj.ro/~darvay/eme/mtne2017/">(https://www.cs.ubbcluj.ro/~darvay/eme/mtne2017/)</a></p>    | D | 1 |
| <p>Membru comitet organizare + comitet științific: A Magyar Tudomány Napja Erdélyben, 7. Matematika és Informatika Alkalmazásokkal Konferencia, 2016. november 18-20.<br/> <a href="https://www.cs.ubbcluj.ro/~darvay/eme/mtne2016/">(https://www.cs.ubbcluj.ro/~darvay/eme/mtne2016/)</a></p>  | D | 1 |

|  |   |   |
|--|---|---|
| Membru în comitetul de organizare a conferinței MACS 2014 ( <a href="https://www.cs.ubbcluj.ro/~macs/2014/">https://www.cs.ubbcluj.ro/~macs/2014/</a> )  | D | 1 |
| Membru în comitetul de organizare a școlii de vară "Domain specific languages summer school", 2013, UBB, Cluj-Napoca ( <a href="http://dsl2013.math.ubbcluj.ro/organizers.html">http://dsl2013.math.ubbcluj.ro/organizers.html</a> )<br>( <a href="https://link.springer.com/book/10.1007/978-3-319-15940-9">https://link.springer.com/book/10.1007/978-3-319-15940-9</a> , LNCS, volume 8606) | C | 2 |
| Membru în comitetul de organizare a școlii de vară "Central European Functional Programming School 2007", UBB, Cluj-Napoca ( <a href="http://www.cs.ubbcluj.ro/cefp2007/">http://www.cs.ubbcluj.ro/cefp2007/</a> )<br>( <a href="https://link.springer.com/book/10.1007/978-3-540-88059-2">https://link.springer.com/book/10.1007/978-3-540-88059-2</a> , LNCS 5161)                           | C | 2 |
| <b>Keynote/Invited speaker: 6</b>  |   |   |
| 9th International Conference on Mathematics and Informatics, September 7-8, 2023, Târgu Mureș/Marosvásárhely, Romania ( <a href="https://mathinfo.ms.sapientia.ro/2023/">https://mathinfo.ms.sapientia.ro/2023/</a> ,<br><a href="https://mathinfo.ms.sapientia.ro/2023/Program_Mathinfo_Info.pdf">https://mathinfo.ms.sapientia.ro/2023/Program_Mathinfo_Info.pdf</a> )                       | D | 1 |
| Ifjú Székely Informatikusok Kollokviuma (ISZIK) 2021 ( <a href="http://locsi.web.elte.hu/masz-im/iszik2021-program.pdf">http://locsi.web.elte.hu/masz-im/iszik2021-program.pdf</a> )   | D | 1 |
| Pannonian Conference on Advances in Information Technology (PCIT) 2019 ( <a href="https://pcit2019.mik.uni-pannon.hu/index.php/en/">https://pcit2019.mik.uni-pannon.hu/index.php/en/</a> )   | D | 1 |
| 7. Digitális Székelyföld Konferencia 2019 ( <a href="https://digitalisszekelyfold.ro/7-digitalis-szekelyfold-konferencia">https://digitalisszekelyfold.ro/7-digitalis-szekelyfold-konferencia</a> )  | D | 1 |
| Conferința națională Magyar Tudomány Napja Erdélyben, 25-26. November, 2016, Cluj-Napoca: "Költséghatékony osztályozás: a félig felügyelt gépi tanulási algoritmusok" ( <a href="http://www.eme.ro/mtne">http://www.eme.ro/mtne</a> )  | D | 1 |

|  |   |   |
|--|---|---|
| <p>Conferința internațională 11<sup>th</sup> Joint Conference on Mathematics and Computer Science, 20-22 May, 2016, Eger, Hungary: "Similarity and kernels in machine learning". (<a href="http://macs.elte.hu/">http://macs.elte.hu/</a>, <a href="http://www.cs.ubbcluj.ro/~zbodo/pres/macs2016.pdf">http://www.cs.ubbcluj.ro/~zbodo/pres/macs2016.pdf</a>)</p>  | D | 1 |
| <b>Profesor/Cercetător invitat/asociat</b>   |   |   |
|  |   |   |
| <b>Consolidarea de echipe de cercetare</b>   |   |   |
|  |   |   |
| <b>Membru in comisii de evaluare a tezelor de doctorat la universitate de top</b>  |   |   |
|  |   |   |
| <b>Membru în comisii de îndrumare a doctoranzilor</b>  |   |   |
|  |   |   |
| <b>Dezvoltarea de pachete software: 2</b>  |   |   |
| <p>Porter Stemmer în GNU Flex: <a href="https://tartarus.org/martin/PorterStemmer/">https://tartarus.org/martin/PorterStemmer/</a> + <a href="http://www.cs.ubbcluj.ro/~zbodo/?a=soft-data/porter_flex">http://www.cs.ubbcluj.ro/~zbodo/?a=soft-data/porter_flex</a></p> <p>Martin Porter's website (traffic: <a href="https://www.similarweb.com/website/tartarus.org/#overview">https://www.similarweb.com/website/tartarus.org/#overview</a>)</p> |   | 2 |
| <b>Pozitii de conducere în organizatii profesionale:</b>   |   |   |

|   |  |   |
|---|--|---|
|   |  |   |
| <b>Premii si alte merite: 4</b>   |  |   |
| PREMIUL PENTRU EXCELENȚĂ ȘTIINȚIFICĂ, Kolozsvári Akadémiai Bizottság / FILIALA DIN CLUJ A ACADEMIEI MAGHIARE DE ȘTIINȚE, 2022   |  | 2 |
| Premiul "Active Learning Challenge Award", primul loc în clasificarea documentelor (setul D) Active Learning and Experimental Design Workshop, May 16, 2010, Sardinia, Italy (AISTATS workshop).<br>( <a href="http://clopinet.com/isabelle/Projects/AISTATS2010/">http://clopinet.com/isabelle/Projects/AISTATS2010/</a> , <a href="http://www.causality.inf.ethz.ch/activelearning.php">http://www.causality.inf.ethz.ch/activelearning.php</a> ) |  | 2 |
| <b>Brevete și invenții active</b>   |  |   |
|   |  |   |

**Criteriu director proiect de cercetare cu minim 2 membri, obținut prin competiție la nivel național sau internațional:**

Director proiect de cercetare finanțat de MTA (**Academia Maghiară de Științe**) Domus, *Magyar könnyűzenei dalszövegek elemzése statisztikai természetesnyelv-feldolgozási módszerekkel (Analyzing Hungarian popular music lyrics using statistical natural language processing methods)*, grant ID 91/22/2023/HTMT

- sumă grant: 400000 Ft
- perioada de desfășurare: 1 aprilie 2023 - 15 noiembrie 2023
- membrii proiectului:
  - coordonator proiect: conf. dr. Zalán-Péter Bodó

- membri: asist. dr. Annamária Szenkovits, drd. Attila Mester
- pachetul de informații (în lb. maghiară): <https://mta.hu/domus-osztondij/a-domus-osztondij-105505>,  
[https://mta.hu/data/dokumentumok/magyar\\_tudomanyossag\\_kulfoldon/2023/Domus\\_szulofoldi\\_2\\_osztondij\\_palyazati\\_felhivas\\_2023.pdf](https://mta.hu/data/dokumentumok/magyar_tudomanyossag_kulfoldon/2023/Domus_szulofoldi_2_osztondij_palyazati_felhivas_2023.pdf)

Anexe:

- Contract de cercetare (în lb. maghiară)
- Propunerea de cercetare (în lb. maghiară) - din care reiese că componența echipei este prestabilit în aplicație, fiecare membru având propriul său subtemă. Activitatea științifică a membrilor a fost, de asemenea, luată în considerare la evaluarea aplicației.

## ÖSZTÖNDÍJ SZERZŐDÉS

amely létrejött egyrészről

**a Magyar Tudományos Akadémia Titkársága** (székhely: 1051 Budapest, Nádor utca 7., törzsszám: 300201; Áht-azonosító szám: 039398; adószám: 15300203-2-41), mint a támogatás folyósításával kapcsolatos feladatok lebonyolításával és a szakmai koordinációval kapcsolatos feladatok ellátására kijelölt költségvetési szerv, támogató, (a továbbiakban **Támogató**), képviseli: meghatalmazás alapján Oberfrank Ferenc, igazgató, **Köztestületi Igazgatóság**,

másrészről

hivatalos név: Bodó Zalán Péter

születési név: Bodó Zalán Péter

anyja neve: Debreczeni Erzsébet

születési hely, idő: Szatmárnémeti, 1981.02.16

állampolgárság:

állandó lakcím: 407505, Tordaszentlászló, 237, Románia

telefonszám: +40 741 706839

személyi igazolvány/útlevélszám: CJ 642228

magyarországi/szülőföldi adószám (amennyiben van ilyen):

magyarországi TAJ szám (amennyiben van ilyen):

bankszámla tulajdonosának neve: Zalán Péter Bodó

számlavezető bank megnevezése: Banca Transilvania

számlavezető bank címe: Calea Dorobantilor 30-36, 400117, Cluj-Napoca, Romania

számlaszám: RO92BTRLHUF CRT0012702201

IBAN kód: RO92BTRLHUF CRT0012702201

swiftkód: BTRLRO22

utalható devizanem: HUF

továbbiakban **Ösztöndíjas** - együttesen: Felek - között, az alulírott napon és helyen az alábbi feltételekkel:

### Előzmény

1. A Magyar Tudományos Akadémia (a továbbiakban: MTA) Domus szülőföldi ösztöndíj elnyerésére pályázati felhívást tett közzé, amellyel támogatni kívánja a Kárpát-medencei kutatók szülőföldi munkáját, a magyar tudományos eredmények létrehozását, a meglévő magyar szellemi értékek dokumentálását; a tudományos műhelyek erősítését, a magyar tudományos szervezetek együttműködését, Kárpát-medencei magyar tudományos hálózatok létrejöttét és előkészítését; a tudományos utánpótlás segítését, korszerű oktatási segédanyagok kidolgozását, színvonalas magyar tankönyvek megjelentetését, Kárpát-medencei szintű doktori iskolák létrehozását, annak előkészítését.

Ennek érdekében az MTA elnöke a magyar tudományosság határon túli gondozására, az akadémiai támogatások gazdálkodásával kapcsolatos feladatok lebonyolítására 2012. július 1-jei hatállyal a Támogatót jelölte ki. Az ösztöndíjak pénzügyi fedezetét az „MTA Titkárság Igazgatása, tiszteletdíjak” költségvetési cím tartalmazza. A szerződés szakmai adminisztrációja az MTA Titkárság Határon Túli Magyarok Titkársága (a továbbiakban: Határon Túli Magyarok Titkársága) feladata.

### Az ösztöndíj célja

2. Az MTA hagyományosan támogatja a környező országokban működő magyar tudományos műhelyeket, kapcsolatot tart a külföldön élő, magyar nyelvű és tárgyú tudományos kutatások művelőivel, támogatja a határon túli magyar tudományosságot. Az MTA Domus ösztöndíjprogramja segíti a szülőföldi magyar nyelvű és tárgyú kutatások folytatását, a Kárpát-medencei magyar kutatók együttműködését.
3. A beérkezett pályázatok alapján a Domus szülőföldi ösztöndíjak odaítéléséről az MTA Domus Kuratóriuma döntött. Az MTA Domus Kuratórium döntése alapján csoportos kutatási ösztöndíj jogcímen **400 000 Ft**, azaz **négyszázezer forint** ösztöndíjat ítél meg. Ösztöndíjas részére a Támogatóhoz benyújtott pályázati dokumentáció részét képező kutatási terv megvalósítására.
4. Az ösztöndíjas tevékenység időtartama: **2023. április 1. - 2023. november 15.**
5. Az Ösztöndíjas az alábbi tudományos tevékenység folytatásához részesül támogatásban:  
*Magyar könnyűzenei dalszövegek elemzése statisztikai természetesnyelv-feldolgozási módszerekkel.*

### Az ösztöndíj mértéke és feltételei

6. Támogató kötelezettséget vállal arra, hogy a 3. pontban megjelölt mértékű ösztöndíjat **két egyenlő részletben** folyósítja az Ösztöndíjas részére, az alábbi ütemezésben:
  - az ösztöndíj I. részletének, **200 000** forintnak az átutalására a jelen szerződés aláírásától számított 30 napon belül kerül sor;
  - az ösztöndíj II. részletének, **200 000** forintnak az átutalására az Ösztöndíjas által feltöltött szakmai beszámoló és pályamunka Domus Kuratórium általi elfogadását követő 5 napon belül kerül sor.
7. A támogatás hazai költségvetési forrásból finanszírozott nemzetközi

oktatási, képzési, kutatásfejlesztési és kulturális mobilitási programok keretében belföldi és külföldi magánszemélyek részére folyósított ösztöndíjnak minősül, így az elszámolt költségek az 1995. évi CXVII. törvény (Szja törvény) 1. számú melléklete 4.7. pont c) alpontja értelmében adómentesek.

### **Az Ösztöndíjas beszámolási és további kötelezettségei**

8. Az Ösztöndíjas a benyújtott kutatási tervben leírt kutatómunkájáról, annak eredményéről szakmai beszámolót, valamint költségelszámolást készít, amelyet a pályamunkával együtt köteles az Ösztöndíjas időszak lezárásáig (2023. november 15.) feltölteni a Domus elektronikus pályázati rendszerbe, a rendszerből letölthető beszámoló sablon formátumban és a pályamunka fedőlap használatával.
9. A Támogató jogosult a szakmai beszámoló és pályamunka kiegészítését kérni a benyújtást követő 15 napon belül, ha az abban foglaltak nem igazolják a szerződés megfelelő teljesítését. Az Ösztöndíjas a pályamunka kiegészítésére irányuló felhívásnak 30 napon belül köteles eleget tenni. Ha a Támogató a pályamunka kiegészítése iránti felhívással nem él, a benyújtott pályamunkát elfogadottnak tekinti.
10. Az MTA pénzügyi támogatásával létrejövő tudományos mű hosszú távú megőrzését, annak a tudományos közösség számára történő hozzáférhetővé tételét támogatandó, az Ösztöndíjas hozzájárul az átadott, és szakmailag elbírált pályamunkának az MTA Könyvtár és Információs Központ Kézirattárában történő elhelyezéséhez.
11. Amennyiben az Ösztöndíjas a Támogató által támogatott kutatási témában publikációt jelentet meg, vagy ezzel kapcsolatban nyilvános előadást tart, hivatkozni kell arra, hogy a mű, a kutatási eredmény az MTA Domus Programja által támogatott tevékenység keretében jött létre. Az Ösztöndíjas a megjelent publikációból, előadás kéziratából két tiszteletpéldányt köteles eljuttatni az MTA Titkárság Határon Túli Magyarok Titkárságának. Az ösztöndíjasok megjelent munkáit a Domus 'Collegium Hungaricum' könyvgyűjteménye fogadja be és őrzi, valamint teszi elérhetővé az érdeklődők számára.

### **A szerződés megszűnése**

12. Felek megállapodnak abban, hogy a jelen szerződésben, a vonatkozó jogszabályokban, avagy a Támogató által kiírt pályázati felhívásban meghatározottaktól eltérő tevékenység, magatartás vagy nyilatkozat, illetve a beszámolási, és költség-elszámolási kötelezettség késedelmes teljesítése szerződésszegésnek minősül, amely esetben Támogató a jelen szerződést felmondhatja, az Ösztöndíjast meghatározott időre kizárhatja a későbbiekben kiírandó pályázatokból és jogosult a támogatás teljes vagy részbeni visszakövetelésére.
13. Amennyiben az Ösztöndíjas jelen szerződésben foglalt kötelezettségének a Támogató írásbeli felszólításának kézhezvételétől számított 30 napon

belül nem tesz eleget, Támogató a jelen szerződést azonnali hatállyal felmondhatja, ebben az esetben Ösztöndíjas köteles a Polgári Törvénykönyvről szóló 2013. évi V. törvény (Ptk.) szerinti kamattal, késedelem esetén késedelmi kamattal növelt mértékben a támogatás összegét visszafizetni Támogató részére. A támogatás felhasználásával összefüggő, vagy a jelen szerződés szerinti további támogatotti kötelezettségek megszegése esetén egyebekben a hatályos államháztartási jogszabályok rendelkezései irányadók.

14. A jelen szerződésben rögzített bármely jogkövetkezmény tényleges alkalmazásáról a Támogató az eset összes körülményeinek mérlegelésével dönt.

#### **Adatvédelem**

15. Az Ösztöndíjas a jelen szerződésben foglaltak teljesítése céljából aláírásával tudomásul veszi, hogy Támogató a személyes adatait kezeli. Támogató tájékoztatja Ösztöndíjast, hogy adatkezelési tájékoztatója a [www.mta.hu](http://www.mta.hu) internetes oldallap alján a „Hasznos információk” cím alatt tekinthető meg.

#### **Bírósági illetékesség**

16. Felek megállapodnak abban, hogy a jelen szerződésre a magyar anyagi és eljárési jog szabályait rendelik alkalmazni.
17. Felek a jelen szerződéssel kapcsolatos vitás kérdéseiket elsődlegesen tárgyalásos úton, egyezségkötés útján kötelesek rendezni. Amennyiben a jogvita békés úton történő rendezése nem vezet eredményre, úgy a Felek kikötik a Budai Központi Kerületi Bíróság kizárólagos illetékességét.

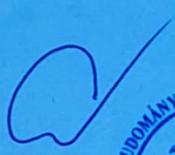
#### **Általános és záró rendelkezések**

18. Ösztöndíjas a jelen szerződés aláírásával igazolja, hogy az ösztöndíjat elfogadja.
19. Ösztöndíjas tudomásul veszi, hogy a szerződésben foglaltakkal összefüggésben felmerülő károkért a magyar polgári jog szabályai szerint tartozik felelősséggel a Támogató felé.
20. A jelen szerződést érintő bármely módosításról a Felek írásban kötelesek megállapodni. Az így létrejövő szerződésmódosítást csatolni kell az eredeti szerződéshez.
21. Felek a jelen szerződést, mint akaratukkal mindenben megegyezőt, elolvasás és értelmezés után jóváhagyólag írták alá azzal, hogy a szerződésben nem szabályozott kérdések tekintetében a vonatkozó hatályos jogszabályok, így különösen a Polgári Törvénykönyvről szóló 2013. évi V. törvény, az államháztartásról szóló 2011. évi CXCV.

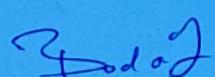
törvény, az államháztartásról szóló törvény végrehajtásáról szóló 368/2011. (XII.31.) Korm. rendelet és a szerzői jogról szóló 1999. évi LXXVI. törvény rendelkezései az irányadók.

22. Jelen ösztöndíj szerződés három eredeti, egymással teljes egészében megegyező példányban készült, amelyből aláírás után két példány a Támogatónál és egy példány az Ösztöndíjasnál marad.

Budapest, 2023. április hó 21. nap

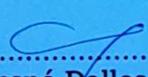
  
.....  
Oberfrank Ferenc  
igazgató  
MTA Titkársága  
Köztestületi Igazgatóság  
Támogató képviselőként



  
.....  
Ösztöndíjas

Pénzügyi ellenjegyző:

2023 ÁPR 21.

  
.....  
Kondácsné Dallos Sarolta  
gazdasági igazgató  
MTA Gazdasági Igazgatóság



*Domus szülőföldi ösztöndíj  
2023*

*Bodó Zalán*

*Szenkovits Annamária  
Mester Attila*

*Magyar könnyűzenei dalszövegek elemzése statisztikai  
természetesnyelv-feldolgozási módszerekkel*

## CSOPORTOS KUTATÁSI SZÁNDÉKNYILATKOZAT

### *Magyar könnyűzenei dalszövegek elemzése statisztikai természetesnyelv-feldolgozási módszerekkel*

**Csoportvezető neve, országa:** Bodó Zalán, Románia

**Résztvevők neve, országa, résztémája:**

Szenkovits Annamária, Románia, Természetesnyelv-feldolgozási módszerek neurális hálók segítségével

Mester Attila, Románia, Témamodelllezési (topikmodelllezési) módszerek

Az együttműködő felek elfogadják a következő feltételeket és alapelveket:

1. Az együttműködés önkéntes.
2. Az együttműködők vállalják, hogy a kutatás keretében végzett munkáról kölcsönösen tájékoztatják egymást.
3. Az együttműködés keretében gyűjtött tárgyi és dokumentációs anyag a csoportvezető intézményében, vagy a csoport által elfogadott kutatóhelyen marad. Az együttműködésben résztvevők vállalják, hogy a feltüntetett kutatások anyagát megőrzik, igény szerint kutathatóságát biztosítják.
4. Az együttműködésben való részvétel a szerződésben meghatározott időre szól.

Kelt:

2023. március 10.

Pályázó (csoportvezető) aláírása:

Résztvevők aláírása:

## KUTATÁSI TERV

### *Magyar könnyűzenei dalszövegek elemzése statisztikai természetesnyelv-feldolgozási módszerekkel*

**Pályázó neve: Bodó Zalán**

**Ország: Románia**

#### **1. A témaválasztás indokai, előzmények, eddigi eredmények**

Figyelembe véve az utóbbi évek kiemelkedő fejlődését és eredményeit a gépi tanulásban, nem is beszélve a természetesnyelv-feldolgozásról, a nagy nyelvi modellek (LLM) kétségtelenül a mesterséges intelligencia kutatásának leghasznosabb és legsokoldalúbb megvalósításai közé tartoznak. Bár számos példát lehetne ezt alátámasztandó felsorakoztatni, mi csak a legfontosabbak közül emelünk ki kettőt: az emberi pontosságot meghaladó 2018-as kérdésmegválaszolási eredményeket (<https://rajpurkar.github.io/SQuAD-explorer/>), valamint az OpenAI híres ChatGPT-jét (<https://openai.com/blog/chatgpt>), amely 2022 novemberi megjelenése óta felforgatta a közösségi médiát és a híroldalakat, valamint a tudományos világot is.

A kutatások azt mutatják, hogy a nyomtatott – vagy az ezzel ekvivalens digitális/elektronikus – szöveg a leghatékonyabb médium az információ tárolására és átadására, azaz az információ felidézése olvasáskor a legpontosabb, szemben a hang- vagy videofolyamok használatával [7]. Ezért a szöveges adatok bányászata – továbbra is – központi jelentőségű problémának számít, és számos területen alkalmazható, a spamszűréstől a piaci intelligenciáig [17].

A természetes nyelvi megértést MI-teljes problémának tekintjük, azaz e probléma megoldása egyenértékű a teljes mesterséges intelligencia problémájának megoldásával [18]. A nyelvnek azonban megvannak a maga korlátai: sokszor nem egyértelmű, „nagyon specifikus és mélyen korlátozott tudásreprezentáció” [5]. Nem meríti ki a tudást, az emberi tudás nagy része nem nyelvi, „egy pusztán a nyelvre betanított rendszer soha nem fogja megközelíteni az emberi intelligenciát, még akkor sem, ha mostantól az univerzum hőhaláláig tanítják azt” [5]. Ezért az olyan nagy nyelvi modellek, mint a GPT-3 [4] vagy a LaMDA [15] csak sekélyes nyelvi megértéssel fognak rendelkezni – ami természetesen egyáltalán nem von le semmit e rendszerek érdemeiből, illetve nem kérdőjelezi meg e modellek hasznosságát.

Ebben a kutatásban statisztikai természetesnyelv-feldolgozási módszerekkel elemezzük magyar könnyűzenei dalok szövegeit, a célunk a zenei műfajok különböző szempontok szerinti összehasonlítása, a műfajokkal kapcsolatos állítások megerősítése vagy cáfolása (például, hogy a metál zenei műfaj negatív érzelmi polaritású, gyakran tartalmaz szomorúsággal, depresszióval, elszigeteltséggel kapcsolatos szöveges témákat [14]), illetve a kapott eredmények hasonló, angol nyelvű kísérletek kimenetével való összevetése [6,13]. Ennek érdekében szövegosztályozási módszereket alkalmazunk a műfajok hasonlóságának, megkülönböztethetőségének vizsgálatára [6], témamodellézést alkalmazunk a zenei műfajok jellemzésére [9], illetve megvizsgáljuk a dalszövegek komplexitását [16] és hangulatát [12]. A kísérletek elvégzéséhez szükséges egy statisztikailag reprezentatív, magyar nyelvű dalszöveg-adatbázis létrehozása, mivel ilyen – többnyire szerzői jogi problémák miatt – nem áll a rendelkezésünkre.

Eddigi eredményként elmondható, hogy már több kísérletet is sikerült elvégeznünk az eddig feldolgozott adatokon, s ezek két konferencián is bemutatásra kerültek:

1. Szenkovits Annamária: *Pop, rock, esetleg punk vagy folk? Magyar dalszövegek elemzése MI módszerekkel*. 8. Digitális Székelyföld Konferencia, 2022. október 27–29., Csíkszereda  
(<https://digitalisszekelyfold.ro/schedule/kutatas-fejlesztes-mi-a-muveszetekben>)
2. Szenkovits Annamária, Bodó Zsolt: *Pop, rock, esetleg punk vagy folk? Magyar dalszövegek elemzése klasszikus és modern statisztikai és mesterséges intelligenciai módszerekkel*. 2022. november 25–27., Kolozsvár  
(<http://www.cs.ubbcluj.ro/~darvay/eme/mtne2022/program/program2022.pdf>)

Ennek ellenére publikáció még nem született a kutatásból.

A csoportvezető már több romániai és nemzetközi kutatási projektben vett részt, ahol természetesnyelv-feldolgozási problémákkal is foglalkozott. A jelen kutatáshoz tematikában legközelebb álló munka a [3], amiben hasonlóképpen egy dalszöveg-adatbázis került megépítésre, melyen különböző zeneműfaji osztályozásokat végeztünk. Az automatikus osztályozás megvalósításához szöveges n-gramokat, rím alapú jellemzőket, illetve szövegstatistikai jellemzőket alkalmaztunk. További kapcsolódó munkáknak tekinthetők a [1,2,8,10,11] publikációk.

## 2. A kutatás célja

A kutatás célja a zenei műfajok különböző szempontok szerinti összehasonlítása, a műfajokkal kapcsolatos, a szakirodalomban megjelenő állítások megerősítése vagy cáfolása (például, hogy a metál zenei műfaj negatív érzelmi polaritású, gyakran tartalmaz szomorúsággal, depresszióval, elszigeteltséggel kapcsolatos szöveges témákat [14]), illetve a kapott eredmények hasonló, angol nyelvű kísérletek kimenetével való összevetése [6,13]. A kísérletek elvégzéséhez szükséges egy statisztikailag reprezentatív, magyar nyelvű dalszöveg-adatbázis létrehozása, mivel ilyen – többnyire szerzői jogi problémák miatt – nem áll a rendelkezésünkre. Habár a dalszöveg-adatbázis fontos részét képezi a kutatási munkafolyamatnak, a fent említett problémák miatt ez nem tehető közzé, viszont a kutatás részeként tervezzük egy a kutatást részletesen bemutató honlap létrehozását is, ahol minden statisztikát, grafikont, eredményt publikálni kívánunk a további hasonló irányú kutatások elősegítése érdekében.

## 3. A kutatás módszertana

A jelen kutatás módszertana illeszkedik az adatbányászati kutatások általában követett metodológiájára:

A probléma, célok megfogalmazása (ez részben már megtörtént).

Hipotézisek megfogalmazása.

Adatgyűjtés (ez részben már megtörtént).

Módszerek tanulmányozása, implementálása (ez részben már megtörtént).

Kísérletek elvégzése, az eredmények elemzése.

Az eredmények publikálása.

## 4. A kutatás időbeli ütemezése

2023. április: Adatbázis frissítése, azaz a létező adatbázis kiegészítése új dalszövegekkel. Jelen pillanatban egy 2021. decemberében letöltött dalszöveg-adatbázissal dolgozunk.

2023. május: Adatbázis megtisztítása (üres dalszövegek eltávolítása, nyelvi szűrés), hogy az a kísérletek elvégzéséhez szükséges formában legyen. Az adatbázis összekapcsolása a MusicBrainz ([https://musicbrainz.org/doc/MusicBrainz\\_API](https://musicbrainz.org/doc/MusicBrainz_API)) és Last.fm (<https://www.last.fm/api>) szolgáltatások által visszakapott metainformációkkal a műfajok hozzárendelése végett. Az adatok műfaj szerinti szűrése.

2023. június–július: A kísérletek elvégzése: dalszövegek műfaji osztályozása, témamodellezés, szövegbonyolultság, hanguletelemzés.

2023. augusztus–szeptember: A tanulmány megírása, illetve konferenciára/szakfolyóiratba való beküldése. A weboldal létrehozása, ahol a kísérletek részletes eredményeit publikáljuk.

## 5. Költségterv

A tanulmányt elsősorban szakfolyóiratban tervezzük publikálni, a számításba jövő két szakfolyóirat pedig a következő: AI & Society (Springer, <https://www.springer.com/journal/146/>), Speech Communication (Elsevier, <https://www.sciencedirect.com/journal/speech-communication>). A pályázott összeg legnagyobb része a kutatás megvalósításához szükséges személyi költségeket foglalja magába, de tervezzük szakkönyvek, cikkek vásárlását is.

Személyi költségek: 300000 Ft

Cikkek, könyvek vásárlása: 100000 Ft

**Összesen:** 400000 Ft

## 5. A kutatás továbbvitelének lehetőségei

A projekt keretein belül létrejövő kutatási eredményeket fontosnak tartjuk, mivel ilyen jellegű kísérletek magyar nyelvű dalszövegekre korábban nem történtek, míg például angol nyelvű hasonló kutatások nagy számban fellelhetők. Fontos lehet tehát a kapott eredmények összevetése más nyelven történt/történő hasonló eredményekkel, az egyezések, illetve különbségek magyarázata szociális és kulturális szempontból. A létrejövő megtisztított, illetve megfelelő módon feldolgozott és tárolt adatbázis – amit a szerzői jogok miatt nem publikálhatunk, azonban természetesen egyéni kérésre a kutatók rendelkezésére bocsátunk – további kutatások alapját képezheti.

## 6. A kutatás rövid leírása angol nyelven

Considering the outstanding progress and achievements of the last few years in machine learning, let alone natural language processing, large language models are undeniably among the most useful and versatile realizations of artificial intelligence research. Although many examples could be given here, we pick just two of the most important ones: question answering results beyond human accuracy on the SQuAD2.0 dataset (<https://rajpurkar.github.io/SQuAD-explorer/>), and OpenAI's famous ChatGPT (<https://openai.com/blog/chatgpt/>) that has been all over social media and news sites since its publication in November 2022.

Research shows that printed – or equivalently digital – text is the most efficient medium for information storage and transfer, i.e. recalling information is most accurate when reading, compared to using audio or video streams [7]. Therefore, mining textual data is – still – considered a problem of central importance, having applications in various fields, from spam filtering to market intelligence [17].

In this research, we use statistical natural language processing methods to analyze the lyrics of Hungarian popular music songs, the objective being to compare music genres from different perspectives, confirming – or refuting – claims related to genres (such as metal music is of negative sentiment polarity, often containing lyrical themes of sadness, depression, isolation, etc. [14]) or comparing the obtained results to that of similar, English language experiments [6,13]. To this end, we use text classification methods to examine class similarities, apply topic modeling to characterize musical genres, and investigate the complexity and sentiment of the lyrics. In order to carry out the experiments, it is necessary to create a statistically representative database of Hungarian-language lyrics, as such a database is not available, mostly due to copyright issues. Although creating a song lyrics database is an important part of the research process, it cannot be published due to the problems mentioned above, but as part of the research we plan to create a website with a detailed description of the research, where all statistics, graphs and results will be published to facilitate further research in this field.

### *Hivatkozások*

- [1] Zalán Bodó. Fake news detection without external knowledge. In MDIS, volume 1341 of Communications in Computer and Information Science, pages 202–221. Springer, 2021.
- [2] Zalán Bodó, Zsolt Minier, and Lehel Csató. Text categorization experiments using Wikipedia, 2007.
- [3] Zalán Bodó and Eszter Szilágyi. Connecting the Last.fm dataset to LyricWikia and MusicBrainz. Lyrics-based experiments in genre classification. Acta Universitatis Sapientiae, Informatica, 10(2):158–182, 2018.
- [4] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D. Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. In NIPS, volume 33, pages 1877–1901, 2020.
- [5] Jacob Browning and Yann LeCun. AI And The Limits Of Language. Noema, August 23, 2022. <https://www.noemamag.com/ai-and-the-limits-of-language/>.
- [6] Michael Fell and Caroline Sporleder. Lyrics-based analysis and classification of music. In COLING, pages 620–631, 2014.
- [7] Adrian Furnham. Remembering stories as a function of the medium of presentation. Psychological Reports, 89(3):483–486, 2001.
- [8] Mihály Gencsi, Zalán Bodó, and Annamária Szenkovits. Compilation and validation of a large fake news dataset in Hungarian. In SISY, pages 125–130, 2021.
- [9] Florian Kleedorfer, Peter Knees, and Tim Pohle. Oh Oh Oh Whoah! Towards Automatic Topic Detection In Song Lyrics. In ISMIR, pages 287–292, 2008.
- [10] Zsolt Minier, Zalán Bodó, and Lehel Csató. Segmentation-based feature selection for text categorization. In ICCP, pages 53–59, 2006.
- [11] Zsolt Minier, Zalán Bodó, and Lehel Csató. Wikipedia-based kernels for text categorization. In SYNASC, pages 157–164, 2007.

- [12] Kathleen Napier and Lior Shamir. Quantitative sentiment analysis of lyrics in popular music. *Journal of Popular Music Studies*, 30(4):161–176, 2018.
- [13] Markus Schedl. Genre differences of song lyrics and artist wikis: an analysis of popularity, length, repetitiveness, and readability. In *The World Wide Web Conference*, pages 3201–3207, 2019.
- [14] Leah Sharman and Genevieve A. Dingle. Extreme metal music and anger processing. *Frontiers in Human Neuroscience*, page 272, 2015.
- [15] Romal Thoppilan, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin, Taylor Bos, Leslie Baker, Yu Du, et al. LaMDA: Language Models for Dialog Applications. *arXiv preprint arXiv:2201.08239*, 2022.
- [16] Michael E.W. Varnum, Jaimie Arona Krems, Colin Morris, Alexandra Wormley, and Igor Grossmann. Why are song lyrics becoming simpler? A time series analysis of lyrical complexity in six decades of American popular music. *PloS one*, 16(1):e0244576, 2021.
- [17] Sholom M. Weiss, Nitin Indurkha, Tong Zhang, and Fred Damerau. *Text mining: predictive methods for analyzing unstructured information*. Springer, 2010.
- [18] Roman V. Yampolskiy. Turing test as a defining feature of AI-completeness. In *Artificial intelligence, evolutionary computing and metaheuristics*, pages 3–17. Springer, 2013.