

NADAV BORENSTEIN

PhD student ~ NLP

 nadav.dk

 nb@di.ku.dk

 +45 50 17 36 75

 /in/nadavbor

 Copenhagen, Denmark

SUMMARY

I am a third year Ph.D. candidate in Computer Science at the University of Copenhagen, where I work on the interpretability of neural language models and low resource domains. I am experienced in both the practical and theoretical aspects of Natural Language Processing and Computer Vision.

KEYWORDS

Concepts: NLP, LLMs, CV, Deep Learning, RL, Low resource, multimodal, XAI, Data Science.

Tools: Python, HuggingFace, Tensorflow, PyTorch, scikit-learn, NLTK, spaCy, wandb, Pandas.

EDUCATION

- 9/2021 - 2/2025 **Ph.D. candidate in Computer Science** University of Copenhagen
Under the supervision of Prof. Isabelle Augenstein
- 9/2018 - 10/2020 **M.Sc. Computer Science and Computational Biology** The Hebrew University of Jerusalem
Under the supervision of Prof. Dafna Shahaf
Thesis: BARcode – A Flexible Learning Framework for Biologically Inspired Design
Honors: Excellence Scholarship for M.Sc. Students in Computer Science; magna cum laude
GPA: 96.5%
- 9/2015 - 9/2018 **B.Sc. Computer Science and Computational Biology** The Hebrew University of Jerusalem
Honors: summa cum laude
GPA: 96

EXPERIENCE

- 10/2023 – 3/2024 **Applied Science Intern** Amazon
Internship
Developed a novel method of personification of Large Language Models for customer simulation and wrote an academic paper summarizing the findings
LLMs / Personification
- 5/2022 – 2/2023 **Teaching assistant** University of Copenhagen
The course "Introduction to Natural Language Processing"
Teaching / NLP
- 10/2020 – 8/2021 **Deep Learning Tech lead** Lightricks
Led the research and development of a company-wide project in the field of recommendation systems, including the management of 1-2 additional researchers and collaboration with other departments in the company.
Tech Lead / Recommendation Systems
- 9/2019 – 10/2020 **Deep Learning researcher** Lightricks
Student Position
Developed neural models for automatically enhancing images using Reinforcement Learning and Generative Adversarial Networks that were integrated into the company's product.
GANs / RL
- 4/2019 – 8/2019 **Teaching assistant** Hebrew University of Jerusalem
The course "Computational Models, Computability and Complexity"
Teaching / Computability
- 6/2017 – 11/2018 **Algorithm developer** Mobileye, an Intel Company
Student Position
Developed and implemented deep neural networks and algorithms in C++ and Python designed to detect road hazards for autonomous driving.
CV / Self Driving Cars

PUBLICATIONS

Conferences

- [1] **N. Borenstein**, P. Rust, D. Elliott, and I. Augenstein. "PHD: Pixel-Based Language Modeling of Historical Documents". In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Ed. by H. Bouamor, J. Pino, and K. Bali. Singapore: Association for Computational Linguistics, Dec. 2023, pp. 87–107. doi: 10.18653/v1/2023.emnlp-main.7. URL: <https://aclanthology.org/2023.emnlp-main.7>.

- [2] **N. Borenstein**, K. Stanczak, T. Rolskov, N. Klein Käfer, N. da Silva Perez, and I. Augenstein. "Measuring Intersectional Biases in Historical Documents". In: *Findings of the Association for Computational Linguistics: ACL 2023*. Ed. by A. Rogers, J. Boyd-Graber, and N. Okazaki. Toronto, Canada: Association for Computational Linguistics, July 2023, pp. 2711–2730. DOI: 10.18653/v1/2023.findings-acl.170. URL: <https://aclanthology.org/2023.findings-acl.170>.
- [3] **N. Borenstein**, N. da Silva Perez, and I. Augenstein. "Multilingual Event Extraction from Historical Newspaper Adverts". In: *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Ed. by A. Rogers, J. Boyd-Graber, and N. Okazaki. Toronto, Canada: Association for Computational Linguistics, July 2023, pp. 10304–10325. DOI: 10.18653/v1/2023.acl-long.574. URL: <https://aclanthology.org/2023.acl-long.574>.
- [4] C. Shani*, **N. Borenstein***, and D. Shahaf. "How Did This Get Funded?! Automatically Identifying Quirky Scientific Achievements". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Ed. by C. Zong, F. Xia, W. Li, and R. Navigli. Online: Association for Computational Linguistics, Aug. 2021, pp. 14–28. DOI: 10.18653/v1/2021.acl-long.2. URL: <https://aclanthology.org/2021.acl-long.2>.
- [5] A. Azulay, T. Halperin, O. Vantzou, **N. Borenstein**, and O. Bibi. "Temporally Stable Video Segmentation Without Video Annotations". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Jan. 2022, pp. 3449–3458. URL: https://openaccess.thecvf.com/content/WACV2022/html/Azulay_Temporally_Stable_Video_Segmentation_Without_Video_Annotations_WACV_2022_paper.html.
- [6] H. Emuna, **N. Borenstein**, X. Qian, H. Kang, J. Chan, A. Kittur, and D. Shahaf. "Imitation of Life: A Search Engine for Biologically Inspired Design". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 38.1 (Mar. 2024), pp. 503–511. DOI: 10.1609/aaai.v38i1.27805. URL: <https://ojs.aaai.org/index.php/AAAI/article/view/27805>.

Preprints

- [7] Y. Wang, R. G. Reddy, Z. M. Mujahid, A. Arora, A. Rubashevskii, J. Geng, O. M. Afzal, L. Pan, **N. Borenstein**, A. Pillai, I. Augenstein, I. Gurevych, and P. Nakov. *Factcheck-GPT: End-to-End Fine-Grained Document-Level Fact-Checking and Correction of LLM Output*. 2023. arXiv: 2311.09000 [cs.CL]. URL: <https://arxiv.org/abs/2311.09000>.
- [8] **N. Borenstein**, A. Arora, L.-A. Kaffee, and I. Augenstein. *Investigating Human Values in Online Communities*. 2024. arXiv: 2402.14177 [cs.SI]. URL: <https://arxiv.org/abs/2402.14177>.

ADDITIONAL ACADEMIC ACTIVITIES

2023	Research visit at Ryan Cotterell's lab, ETH Zurich
2022	Participation in LxMLS-2022 summer school in Lisbon, Portugal
2021-2024	Review of academic papers for WSDM, EMNLP and ACL main tracks
2018-2019	Participation in a student exchange program at the University of Toronto, Canada
2018	Participation in a summer program for learning the Chinese language at East China University of Science and Technology, Shanghai, China
2016	Internship at Prof. Nir Friedman's systems biology lab, the Hebrew University of Jerusalem

SKILLS

- Extensive programming experience with Python, familiarity with JAVA and C++
- Experience working with NLP, image processing, and machine learning algorithms and tools, including HuggingFace, Tensorflow, PyTorch, scikit-learn, NLTK and spaCy
- Excellent autodidactic capabilities, perseverance, high motivation and ambition
- Participated in several large-scale projects both as a teammate and as a project leader
- Experience with managing multiple, competing priorities under tight schedules