NADAV BORENSTEIN, PH.D.

Post-doc ~ NLP

nadav.dk
nb@di.ku.dk
+45 50 17 36 75
/in/nadavbor
Copenhagen, Denmark

SUMMARY

I am a post-doc researcher in Computer Science at the University of Copenhagen, where I work on multimodal fact-checking and misinformation detection. 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.

EXPERIENCE

1/2025 – 4/2027 **Post-doc researcher** — **University of Copenhagen**
Researching multimodal fact-checking and misinformation detection. Funded by the Pioneer Centre for Artificial Intelligence.
`Multimodality / Fact-checking`

10/2023 – 3/2024 Internship **Applied Science Intern** — **Amazon**
Developed a novel method of personification of Large Language Models for customer simulation and wrote an academic paper summarizing the findings.
`LLMs / Personification`

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 Student Position **Deep Learning researcher** — **Lightricks**
Developed neural models for automatically enhancing images using Reinforcement Learning and Generative Adversarial Networks that were integrated into the company's product.
`GANs / RL`

6/2017 – 11/2018 Student Position **Algorithm developer** — **Mobileye, an Intel Company**
Developed and implemented deep neural networks and algorithms in C++ and Python designed to detect road hazards for autonomous driving.
`CV / Self Driving Cars`

EDUCATION

9/2021 - 4/2025 **Ph.D. in Computer Science** — **University of Copenhagen**
Under the supervision of Prof. Isabelle Augenstein
Thesis: Revisiting Noise in Natural Language Processing for Computational Social Science

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%

PUBLICATIONS

Conferences

[1] **N. Borenstein**, G. Warren, D. Elliott, and I. Augenstein. "Can Community Notes Replace Professional Fact-Checkers?" In: *Proceedings of the 63nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 2025. arXiv: `2502.14132 [cs.CL]`. URL: `https://arxiv.org/abs/2502.14132`.

[2] **N. Borenstein**, A. Arora, L.-A. Kaffee, and I. Augenstein. "Investigating Human Values in Online Communities". In: *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*. Ed. by L. Chiruzzo, A. Ritter, and L. Wang. Albuquerque, New

Mexico: Association for Computational Linguistics, Apr. 2025, pp. 1607–1627. ISBN: 979-8-89176-189-6. DOI: 10.18653/v1/2025.naacl-long.77. URL: https://aclanthology.org/2025.naacl-long.77/.

[3] **N. Borenstein**, A. Svete, R. Chan, J. Valvoda, F. Nowak, I. Augenstein, E. Chodroff, and R. Cotterell. "What Languages are Easy to Language-Model? A Perspective from Learning Probabilistic Regular Languages". In: *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Ed. by L.-W. Ku, A. Martins, and V. Srikumar. Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 15115–15134. URL: https://aclanthology.org/2024.acl-long.807.

[4] A. Svete, **N. Borenstein**, M. Zhou, I. Augenstein, and R. Cotterell. "Can Transformers Learn n-gram Language Models?" In: *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. Ed. by Y. Al-Onaizan, M. Bansal, and Y.-N. Chen. Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 9851–9867. DOI: 10.18653/v1/2024.emnlp-main.550. URL: https://aclanthology.org/2024.emnlp-main.550/.

[5] D. Wright, A. Arora, **N. Borenstein**, S. Yadav, S. Belongie, and I. Augenstein. "LLM Tropes: Revealing Fine-Grained Values and Opinions in Large Language Models". In: *Findings of the Association for Computational Linguistics: EMNLP 2024*. Ed. by Y. Al-Onaizan, M. Bansal, and Y.-N. Chen. Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 17085–17112. DOI: 10.18653/v1/2024.findings-emnlp.995. URL: https://aclanthology.org/2024.findings-emnlp.995/.

[6] Y. Wang, R. Gangi Reddy, Z. M. Mujahid, A. Arora, A. Rubashevskii, J. Geng, O. Mohammed Afzal, L. Pan, **N. Borenstein**, A. Pillai, I. Augenstein, I. Gurevych, and P. Nakov. "Factcheck-Bench: Fine-Grained Evaluation Benchmark for Automatic Fact-checkers". In: *Findings of the Association for Computational Linguistics: EMNLP 2024*. Ed. by Y. Al-Onaizan, M. Bansal, and Y.-N. Chen. Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 14199–14230. DOI: 10.18653/v1/2024.findings-emnlp.830. URL: https://aclanthology.org/2024.findings-emnlp.830/.

[7] **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.

[8] **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.

[9] **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.

[10] 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.

[11] A. Azulay, T. Halperin, O. Vantzos, **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.

[12] 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.

Thesis

[13] **N. Borenstein**. "Revisiting Noise in Natural Language Processing for Computational Social Science". Available at https://arxiv.org/abs/2503.07395. PhD thesis. University of Copenhagen, 2025.

TEACHING ACTIVITIES

5/2022 – 2/2023 **Teaching assistant** — **University of Copenhagen**
The course "Introduction to Natural Language Processing".

4/2019 – 8/2019 **Teaching assistant** — **Hebrew University of Jerusalem**
The course "Computational Models, Computability and Complexity".

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