

Bashar Alhafni

Assistant Professor of Natural Language Processing
Mohamed bin Zayed University of Artificial Intelligence (MBZUAI)
bashar.alhafni@mbzuai.ac.ae
<https://basharalhafni.com>

RESEARCH INTERESTS

Areas: Natural Language Processing (NLP), Computational Linguistics, Artificial Intelligence, Machine Learning
Specialty: Arabic NLP, Controlled Generation, Writing Assistants, Educational NLP, Error Detection and Correction, Text Simplification, Readability Assessment, Human-Centered NLP

EDUCATION

New York University Ph.D. in Computer Science Thesis: Controlled Natural Language Generation for Morphologically Rich Languages: The Case of Arabic Advisor: Nizar Habash	New York, NY Aug 2020 – May 2025
University of Southern California M.S. in Computer Science Advisors: Jonathan May , Nanyun Peng	Los Angeles, CA Aug 2017 – May 2019
University of Bridgeport B.S. in Computer Science Minor in Mathematics	Bridgeport, CT Aug 2013 – May 2017

EMPLOYMENT

Mohamed bin Zayed University of Artificial Intelligence Assistant Professor of Natural Language Processing	Abu Dhabi, UAE Aug 2025 – Present
New York University Abu Dhabi Graduate Research Assistant	Abu Dhabi, UAE Sep 2019 – May 2025
Applied Research Scientist Intern Grammarly Mentors: Vipul Raheja , Vivek Kulkarni , Dhruv Kumar	San Francisco, CA May 2023 – April 2024
Research Intern Dataminr Mentors: Joel Tetreault , Ke Zhang , Shihao Ran	New York, NY May 2022 – Sep 2022
Graduate Student Researcher USC Information Sciences Institute	Los Angeles, CA May 2018 – Aug 2019
Associate Software Engineer Pratt & Whitney	East Hartford, CT Nov 2017 – May 2018
Technology Summer Analyst Morgan Stanley	New York, NY Summer 2016 & 2017

PUBLICATIONS

[Google Scholar](#): 984 citations; 10 h-index; 10 i10-index (June, 2025)

- [1] **Bashar Alhafni** and Nizar Habash. 2025. [Enhancing text editing for grammatical error correction: Arabic as a case study](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Vienna, Austria. Association for Computational Linguistics.
- [2] Chatrine Qwaider, **Bashar Alhafni**, Kirill Chirkunov, Nizar Habash, and Ted Briscoe. 2025. [Enhancing Arabic automated essay scoring with synthetic data and error injection](#). In *Proceedings of the 20th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2025)*, Vienna, Austria. Association for Computational Linguistics.

- [3] Kirill Chirkunov, **Bashar Alhafni**, Chatrine Qwaider, Nizar Habash, and Ted Briscoe. 2025. [ARWI: Arabic Write and Improve](#). In *Proceedings of the Fourth Workshop on Intelligent and Interactive Writing Assistants (In2Writing 2025)*, New Mexico, United States. Association for Computational Linguistics. **Diversity Award**.
- [4] **Bashar Alhafni**, Sarah Al-Towaity, Ziyad Fawzy, Fatema Nassar, Fadhil Eryani, Houda Bouamor, and Nizar Habash. 2024. [Exploiting dialect identification in automatic dialectal text normalization](#). In *Proceedings of The Second Arabic Natural Language Processing Conference*, pages 42–54, Bangkok, Thailand. Association for Computational Linguistics.
- [5] Juan Liberato, **Bashar Alhafni**, Muhamed Khalil, and Nizar Habash. 2024. [Strategies for Arabic readability modeling](#). In *Proceedings of The Second Arabic Natural Language Processing Conference*, pages 55–66, Bangkok, Thailand. Association for Computational Linguistics.
- [6] **Bashar Alhafni**, Reem Hazim, Juan David Pineros Liberato, Muhamed Al Khalil, and Nizar Habash. 2024. [The SAMER Arabic text simplification corpus](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 16079–16093, Torino, Italia. ELRA and ICCL.
- [7] Vipul Raheja, Dimitris Alikaniotis, Vivek Kulkarni, **Bashar Alhafni**, and Dhruv Kumar. 2024. [mEdit: Multilingual text editing via instruction tuning](#). In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 979–1001, Mexico City, Mexico. Association for Computational Linguistics.
- [8] **Bashar Alhafni**, Vivek Kulkarni, Dhruv Kumar, and Vipul Raheja. 2024. [Personalized text generation with fine-grained linguistic control](#). In *Proceedings of the 1st Workshop on Personalization of Generative AI Systems (PERSONALIZE 2024)*, pages 88–101, St. Julians, Malta. Association for Computational Linguistics.
- [9] **Bashar Alhafni**, Go Inoue, Christian Khairallah, and Nizar Habash. 2023. [Advancements in Arabic grammatical error detection and correction: An empirical investigation](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 6430–6448, Singapore. Association for Computational Linguistics.
- [10] **Bashar Alhafni**, Ossama Obeid, and Nizar Habash. 2023. [The user-aware Arabic gender rewriter](#). In *Proceedings of the First Workshop on Gender-Inclusive Translation Technologies*, pages 3–11, Tampere, Finland. European Association for Machine Translation.
- [11] Hazem Ibrahim, Fengyuan Liu, Rohail Asim, Balaraju Battu, Sidahmed Benabderrahmane, **Bashar Alhafni**, Wifag Adnan, Tuka Alhanai, Bedoor AlShebli, Riyadh Baghdadi, Jocelyn J. Bélanger, Elena Beretta, Kemal Celik, Moumena Chaqfeh, Mohammed F. Daqaq, Zaynab El Bernoussi, Daryl Fougne, Borja Garcia de Soto, Alberto Gandolfi, and 20 others. 2023. [Perception, performance, and detectability of conversational artificial intelligence across 32 university courses](#). *Scientific Reports*, 13(1).
- [12] **Bashar Alhafni**, Nizar Habash, and Houda Bouamor. 2022. [User-centric gender rewriting](#). In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 618–631, Seattle, United States. Association for Computational Linguistics.
- [13] **Bashar Alhafni**, Nizar Habash, and Houda Bouamor. 2022. [The Arabic parallel gender corpus 2.0: Extensions and analyses](#). In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 1870–1884, Marseille, France. European Language Resources Association.
- [14] **Bashar Alhafni**, Nizar Habash, Houda Bouamor, Ossama Obeid, Sultan Alrowili, Daliyah AlZeer, Kawla Mohamad Shnqiti, Ahmed Elbakry, Muhammad ElNokrashy, Mohamed Gabr, Abderrahmane Issam, Abdelrahim Qaddoumi, Vijay Shanker, and Mahmoud Zyate. 2022. [The shared task on gender rewriting](#). In *Proceedings of the Seventh Arabic Natural Language Processing Workshop (WANLP)*, pages 98–107, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.
- [15] Reem Hazim, Hind Saddiki, **Bashar Alhafni**, Muhamed Al Khalil, and Nizar Habash. 2022. [Arabic word-level readability visualization for assisted text simplification](#). In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 242–249, Abu Dhabi, UAE. Association for Computational Linguistics.
- [16] Hossein Rajaby Faghihi, **Bashar Alhafni**, Ke Zhang, Shihao Ran, Joel Tetreault, and Alejandro Jaimes. 2022. [CrisisLTLSum: A benchmark for local crisis event timeline extraction and summarization](#). In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 5455–5477, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

- [17] Malihe Alikhani, Thomas Kober, **Bashar Alhafni**, Yue Chen, Mert Inan, Elizabeth Nielsen, Shahab Raji, Mark Steedman, and Matthew Stone. 2022. [Zero-shot cross-linguistic learning of event semantics](#). In *Proceedings of the 15th International Conference on Natural Language Generation*, pages 212–224, Waterville, Maine, USA and virtual meeting. Association for Computational Linguistics.
- [18] Go Inoue, **Bashar Alhafni**, Nurpeiis Baimukan, Houda Bouamor, and Nizar Habash. 2021. [The interplay of variant, size, and task type in Arabic pre-trained language models](#). In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, pages 92–104, Kyiv, Ukraine (Virtual). Association for Computational Linguistics.
- [19] **Bashar Alhafni**, Nizar Habash, and Houda Bouamor. 2020. [Gender-aware reinflection using linguistically enhanced neural models](#). In *Proceedings of the Second Workshop on Gender Bias in Natural Language Processing*, pages 139–150, Barcelona, Spain (Online). Association for Computational Linguistics.
- [20] Ossama Obeid, Nasser Zalmout, Salam Khalifa, Dima Taji, Mai Oudah, **Bashar Alhafni**, Go Inoue, Fadhl Eryani, Alexander Erdmann, and Nizar Habash. 2020. [CAMEL tools: An open source python toolkit for Arabic natural language processing](#). In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 7022–7032, Marseille, France. European Language Resources Association.
- [21] Rujun Han, Mengyue Liang, **Bashar Alhafni**, and Nanyun Peng. 2019. [Contextualized word embeddings enhanced event temporal relation extraction for story understanding](#). In *Proceedings of the First Workshop on Narrative Understanding*, Minneapolis, Minnesota. Association for Computational Linguistics.

INVITED TALKS & PRESENTATIONS

- [1] **Controlled Natural Language Generation for Morphologically Rich Languages: The Case of Arabic.** Mohamed bin Zayed University of Artificial Intelligence (MBZUAI). April, 2025.
- [2] **Panelist: Starting and Excelling in Arabic NLP.** Arabic NLP Winter School. January, 2025.
- [3] **Panelist: Arabic Culture and AI: From Preservation to Enrichment.** Arabian Days. December, 2024.
- [4] **Controlled Natural Language Generation for Morphologically Rich Languages: The Case of Arabic.** National Research Council Canada (NRC). September, 2024.
- [5] **Panelist: Challenges and Solutions in Arabic NLP.** King Salman Global Academy for Arabic. September, 2024.
- [6] **User-Centric Gender Rewriting.** Mohamed bin Zayed University of Artificial Intelligence (MBZUAI). May, 2023.
- [7] **Arabic NLP: Challenges and Solutions.** The Symposium on the Evolution of the Arabic Word. October, 2022.
- [8] **Panelist: Young Researchers in Arabic NLP.** The 7th Arabic Natural Language Processing Workshop (WANLP 2022). December, 2022.
- [9] **Gender-Aware Reinflection using Linguistically Enhanced Neural Models.** ETH Zürich. October, 2021.

TEACHING & MENTORING

Undergraduate Research Advising

- **Mohamed Abbas, Abdelrahman Shaheen, Rongxuan Tian, Tereza Truckova** Summer 2025
MBZUAI UGRIP: Tokenization-Free Machine Translation for Dialectal Arabic
- **Sarah Al-Towaity, Ziyad Fawzy, Fatema Nassar** Fall 2023 – Spring 2024
NYU Abu Dhabi Undergraduate Capstone: Dialectal Arabic Text Normalization
- **Juan Piñeros Liberato** Spring 2023 – Spring 2024
NYU Abu Dhabi Undergraduate Capstone: Arabic Readability Assessment
- **Reem Hazim** Spring 2022 – Spring 2024
NYU Abu Dhabi Undergraduate Capstone: Arabic Text Simplification

Courses

- **NYU Abu Dhabi CS-2216: Introduction to NLP** Fall 2024
Taught six lectures on neural networks for NLP
- **NYU Abu Dhabi CS-3090: Research Seminar in Computer Science** Fall 2020
Taught a tutorial on pretrained language models

- **USC CSCI-572: Information Retrieval & Web Search Engines** Fall 2018 – Spring 2019
TA for Ellis Horowitz
- **University of Bridgeport CS-101/102: Intro to Computer Science I & II** Fall 2014 – Spring 2017
TA for Subrina Thompson

Tutorials

- **LLMs in Education: Novel Perspectives, Challenges, and Opportunities** Jan 2025
COLING 2025 Tutorial: [Proposal](#), [Website](#)
- **Arabic NLP with CAMEL Tools** Jan 2025
Tutorial at the Arabic NLP Winter School: [Website](#)
- **Arabic NLP** Jan 2020
Three-day crash course on Arabic NLP for researchers in digital humanities: [Website](#)

PROFESSIONAL SERVICE

Organization Committees

- Co-organizer: Workshop on Innovative Use of NLP for Building Educational Applications (BEA) 2025
- Co-organizer: The BAREC Shared Task on Arabic Readability Assessment at ArabicNLP 2025
- Co-organizer: The Second Arabic Natural Language Processing Conference (ArabicNLP) 2024
- Co-organizer: Gender Rewriting Shared Task at WANLP 2022
- Local chair: EMNLP 2022

Reviewer

- Conferences and Workshops: ACL, EMNLP, NAACL, EACL, AACL, COLING, LREC, BEA, ArabicNLP, GeBNLP
- Journals: Language Resources and Evaluation

Departmental Service

- NLP Department Representative, Teaching & Learning Advisory Board, MBZUAI (2025 – Present)

HONORS & AWARDS

NYU Global PhD Fellowship: 2020 – 2024

Upsilon Pi Epsilon ACM Scholarship: 2017

University of Bridgeport Trustee Scholarship: 2013 – 2017

University of Bridgeport President's List: 2013 – 2017