Mahmoud Zaher

AI Research Engineer

Cairo, Egypt | maredazaher@gmail.com | (+20) 10-6696-4554

zaher-m.me | m-zaher | zaher-m | zaher-m

ABOUT ME

I'm currently an AI research engineer building reliable, trustworthy and explainable AI systems. My recent work spans Arabic NLP, speech, and model interpretability, with three papers accepted at ArabicNLP 2026. I'm always looking for more challenging problems and exploring new research directions.

EDUCATION

BSc in Computer and Control Systems Engineering *Sep 2019 – Jul 2024*

Mansoura University, Egypt | GPA 3.71/4.0

Courses: Computer Architecture, Database Systems, Embedded Systems, Probability and Statistics, Artificial Intelligence, Computer Vision, Machine Learning, Natural Language Processing, Deep Learning

PUBLICATIONS

M Zaher. "ZHR at BAREC Shared Task 2026: Ordinal-Objective Ensembles with Threshold Calibration and Silver-Label Pretraining for Arabic Readability Assessment." *ArabicNLP @ EMNLP*, 2026.

M Zaher, M Alshrief, M Bahgat. "ZHR at StanceEval-2026: Surface-Matched Synthetic Demonstrations and Metric-Aware Decoding for Arabic Cross-Target Stance Detection." *ArabicNLP @ EMNLP*, 2026.

M Alshrief, M Bahgat, **M Zaher**, S Selim. "The Nile at NADI 2026 Subtask 1.1: A Two-Phase Acoustic Fine-Tuning Strategy for Robust Dialectal Arabic Speech Recognition." *ArabicNLP @ EMNLP*, 2026.

WORK EXPERIENCE

AI Research Engineer *Dec 2025 – Present*

Click ITS | Nasr City, Egypt

- Fine-tuned VLMs across different architectures and sizes with LoRA/QLoRA for Arabic financial document extraction, and benchmarked them against NVIDIA NIM models on text, numeric, layout, table, and reading-order fidelity.
- Improved retrieval for colloquial queries by indexing questions predicted from each chunk as separate vectors linked back to it, evaluated in an A/B test with Recall@k, MRR, and nDCG.
- Built the training data stack: a synthetic Arabic financial document generator that uses integer programming for consistent figures and rendered-page label checks, and an annotation platform with model-assisted labeling, reviewer validation, and audit trails.

Full Stack Developer *Jul 2025 – Dec 2025*

Information Technology Institute, Intensive Code Camp | Nasr City, Egypt

- Built full-stack web applications and optimized Django database queries to cut response times.
- Moved long-running and scheduled jobs off the request path with Celery workers, Celery Beat, and Redis.
- Automated testing and deployment through Jenkins and Docker pipelines on Linux.

Data Scientist *Jul 2024 – Dec 2024*

Digital Twin Modeling for Naval Propulsion Systems

- Engineered degradation proxies from propulsion telemetry for mechanical stress, thermal wear, and inefficiency.
- Mitigated multicollinearity with VIF, PCA, and regularization, and benchmarked maintenance cost regressors.

RESEARCH EXPERIENCE

Arabic NLP Shared Tasks *Jul 2026 – Oct 2026*
Independent Researcher – ArabicNLP 2026 @ EMNLP 2026

- Developed a 19-level readability ensemble with five ordinal losses, silver labels, and QWK-tuned thresholds.
- Improved cross-target stance detection with retrieved examples, LoRA-tuned ALLaM-7B, and synthetic data.
- Co-authored a Whisper fine-tuning method for dialectal Arabic ASR that cut WER by 8 points.

Interpretable AI *Aug 2025 – Jan 2026*
Remote Intern – Supervised by a senior PhD researcher, University of Queensland

- Applied gradient-based attribution methods to pre-trained vision and language models.
- Analyzed how saliency maps and weight structures change under post-hoc and intrinsic interpretability methods.
- Studied grokking, where neural networks generalize long after fitting their training data.

Expressive Arabic Text-to-Speech *Oct 2023 – Apr 2024*
Remote Intern – ReachSci STEM Mini-PhD Programme

- Led research on Arabic speech synthesis and reviewed recent generative text-to-speech models.
- Fine-tuned pre-trained GlowTTS and VITS models to improve Arabic synthesis quality.

OPEN-SOURCE CONTRIBUTIONS

DLA Suite *(author, document layout analysis benchmark).* Benchmarked 58 layout configurations across 23 repositories and built a label-free validator for missing text, duplicate regions, and reading-order errors. Repo

AgentScope *(agent framework for large language models).* Fixed failures on empty streamed Anthropic tool calls and duplicate tool call IDs in Ollama. PRs

SynapseKit *(application framework for large language models).* Repaired Mistral chat and embedding integrations after breaking changes in Mistralai client 1.x and 2.x. PRs

OpenProof *(evidence-review system).* Exposed a text recognition failure that attached footnotes to figures by proximity alone. Added a review rule that checks the footnote marker and text first, which the project adopted. Rule

AWARDS

1st Place on unseen targets and **2nd Place** on seen targets in the StanceEval-2026 Arabic Stance Detection Shared Task (22 teams), ArabicNLP 2026, with average F1 of 0.940 and 0.896. Official Results

2nd Place in both sentence-level tracks (restricted and open data) of the BAREC 2026 Arabic Readability Shared Task as the sole author, with quadratic weighted kappa of 85.0 and 85.2 against an 82.5 baseline. Official Results

3rd Place, 2024 Arab IoT & AI Challenge. Qualified to present at GITEX Global 2024, Dubai. Details

Co-founded **ActveX**, a posture-correction fitness app selected for the Incubation Program of Egypt's Ministry of Youth and Sports, in partnership with the Academy of Scientific Research and Technology.

SKILLS

PROGRAMMING	Python, JavaScript, Bash, MATLAB
FRAMEWORKS & TOOLS	PyTorch, Transformers, LangChain, Django, FastAPI, React, Celery, LaTeX
DEVOPS	Docker, Jenkins, Apache, Linux, Git, CI/CD
DATABASES	PostgreSQL, MySQL, MongoDB, Redis, ChromaDB, Neo4j
COMMUNICATION	Cross-team collaboration, stakeholder management, and presenting technical concepts clearly
WRITING	Academic papers, technical milestone reports, and leading R&D proposals
PROJECT MANAGEMENT	Agile and Scrum, planning and coordinating resources, tasks, and people
LANGUAGES	Arabic (native), English (fluent), French (basic)