

AHMET RIFAT OZTURK

AI Researcher · Student Founder · Robotics & Applied AI Builder

Istanbul, Türkiye

ahmetrifat@ethosoft.org · GitHub · LinkedIn · Personal Hugging Face · Ethosoft Hugging Face · ORCID · X · ethosoft.org

High school AI researcher, student founder, and robotics builder working across small language models, Turkish NLP, document and medical AI, autonomous systems, and open-source tooling. First author of a public Research Square preprint, first-listed author of the NEDOQwen technical report, co-founder of Ethosoft, and former VolTrAN Software Captain. Combines research execution, public model and dataset releases, technical leadership, industrial AI work, and hands-on mechanical prototyping.

EDUCATION

Pertevniyal High School · Selective Anatolian High School

Expected Graduation 2028

Entering Grade 11 in the 2026–27 academic year · English preparatory year completed

- Admitted with a 478/500 score on Türkiye's national LGS high school entrance examination.
- Grade averages: 90.46/100 in Grade 9 and 88.81/100 in Grade 10.
- EnglishScore: 458, June 2026.

478 / 500

LGS entrance score

458

EnglishScore

90.46 / 100

Grade 9 average

88.81 / 100

Grade 10 average

Istanbul Technical University · University-Level Academic Exposure

2026

- Audited selected university-level lectures, including Applied Data Engineering, through faculty mentorship from Behçet Uğur Töreyn.
- Observed the installation of high-capacity computing systems and gained exposure to institutional research infrastructure.
- Discussed continuing the TUA lunar-rover navigation project through a future university research opportunity.

RESEARCH & PUBLICATIONS

Foresight Is Not Enough: Sentence-Level Future Signals, Self-Loop Hard Negatives, and the Calibration Gap in Small Language Models

2026

First Author · Primary Research Lead · Research Square preprint, version 1 · DOI 10.21203/rs.3.rs-10331341/v1

Co-authors: Yağız Ekrem Dalar, Ömer Faruk Aksoy, Nedim Mutlu Sezer, Feyzi Arda Salihoğlu · Code and reproducibility repository

- Investigated whether token-level autoregressive language models can acquire sentence-level foresight through an auxiliary future-semantic objective at sentence boundaries.
- Led the idea development, experimental design, training and evaluation pipeline, ablations, reranking analysis, calibration study, reproducibility packaging, and manuscript preparation.
- Ran an approximately 82M-parameter DistilGPT-2-scale study on WikiText-103 and WritingPrompts with three seeds per domain, up to 30,000 training and 2,000 validation sentence pairs per domain.
- Built baseline/core/k-horizon ablations, future-head-aware reranking, domain-transfer evaluation, bootstrap confidence intervals, qualitative analysis, compute accounting, and blind human-evaluation materials.
- Found a stable future-oriented training signal but a behavioral calibration gap: removing the explicit future term changed only 2 of 120 five-candidate selections.
- A lexical self-loop penalty reduced lexical looping from 0.1027 to 0.0546 and semantic looping from 0.0250 to 0.0083.

NEDOQwen: An Auditable Low-Cost Diagnosis-and-Repair Study for a Turkish-Centric 0.824B Language Model

2026

First-Listed Author · Technical-Report Lead · Model-Training Contributor · Technical report prepared

Co-authors: Yağız Ekrem Dalar, Ömer Faruk Aksoy, Nedim Mutlu Sezer, Feyzi Arda Salihoğlu · Code · Base model · Corpus

- Trained and evaluated a custom 824,256,000-parameter Turkish decoder-only language model using a teammate-developed 65,536-token typed-surface tokenizer.
- Architecture: 24 layers, 1,536 hidden size, 16 attention heads, 8 KV heads, 4,096 MLP size, 1,024-token context, RoPE, RMSNorm, grouped-query attention, and SwiGLU.
- Public artifacts include a 60.95B-token uint16 corpus, base and SFT checkpoints, Turkish SFT mixtures, training code, sampling scripts, and documentation.
- On the experimental repair line, TurkishMMLU-sub accuracy rose from 18.11% to 21.00% and TUMLU-mini Turkish from 21.33% to 32.22%.
- Framed the work as a low-cost, auditable diagnosis-and-repair workflow while documenting persistent answer-label bias and deployment limitations.

OPEN-SOURCE RESEARCH & ARTIFACTS

QA-Research-SLM · Evidence-Grounded Research Reasoning2026

Open-Source Developer · GitHub · PEFT adapter · GGUF quantizations

- Developed a Qwen3-1.7B-based research reasoner for structured planning, evidence selection, claim verification, conflict detection, uncertainty calibration, cited synthesis, and strict JSON output.
- Built broad and schema-hard SFT stages followed by verifier-guided rejection sampling and supervised fine-tuning, using Transformers, PEFT, bitsandbytes, QLoRA, and TRUBA/SLURM workflows.
- Released public PEFT and GGUF artifacts with documented diagnostic runs on PubMedQA, HotpotQA, and SciFact.

Relation LM · Anchor-Partner Language Modeling Architecture2026

Open-Source Research Developer · MIT-licensed repository

- Designed an experimental language-model architecture that replaces dense all-pairs attention with explicit anchor-partner relations and routed historical-block selection.
- Released RelationLex, a lossless 16K dual-channel tokenizer separating lexical/punctuation tokens from exact whitespace boundaries.
- Implemented dense and routed PyTorch models, stateful decode, multi-anchor relation blocks, benchmark suites, CPU/CUDA parity tests, and custom Triton selection kernels.
- Reported sparse-vs-dense quality deltas between +0.119% and +0.187% BPB across tested context lengths; packed selection was 8.9–9.3% faster than the generic sparse-selection path in the repository benchmark.

TR-DocVQA-Synth

Dataset Contributor · Public Hugging Face artifact

Contributed to a Turkish Document VQA dataset containing 15,000 document images and 235,000 question-answer pairs across invoices, contracts, offers, orders, and business documents.

Mycelium: Safety Should Be Expensive to Remove

Second Author · Idea Contributor · Manuscript in preparation

Contributed ideas to an AI safety research project studying the cost and behavioral effects of removing safety-bearing structures from language models.

SELECTED TECHNICAL PROJECTS

TUA Lunar Rover Navigation System2026

Python · 3D A* · Soft Actor-Critic · NASA LRO topography · Reinforcement learning

- Developed a lunar-rover navigation stack combining slope-constrained 3D A* global planning with SAC-based local path following.
- Built a custom rover environment with lunar terrain, hazard simulation, lidar-like sensing, pure-pursuit concepts, and 3D visualization.

BioDockX · AI-Agent Bioinformatics Platform2026

FastAPI · React/Vite · RCSB PDB · PubChem · NGL Viewer · Open Babel · AutoDock Vina · Docker

- Built a workflow for protein retrieval, browser-based 3D visualization, ligand creation, receptor/ligand preparation, molecular docking, and structured result reporting.
- Implemented agent tools with deterministic fallback routing and optional Gemini planning for scientific workflow automation.

DataMedX2 · Oncology Clinical AI Orchestration2026

Python · Clinical analytics · AI agents · Decision support · Report generation

- Developed clinical data processing and decision-support workflows for oncology, including patient retrieval, evidence-backed summaries, risk cards, patient timelines, and a doctor-facing panel.
- Integrated specialized sub-agents for clinical analysis and structured reporting in a hackathon production workflow.

VoltranScout

JavaScript · VEX competition analytics

Designed a scouting and match-analysis tool for recording observations, comparing team performance, and supporting alliance and competition strategy.

SLM Stream Chat

JavaScript · Ollama · Hugging Face

Built a local streaming chat UI with model controls, attachments, Modelfile editing, visible reasoning output support, and Ollama/Hugging Face model backends.

EU Code Week Eye-Care Platform

Medical AI · CAD · Mechanical prototyping

Led the mechanical side of a portable AI-assisted eye-care system designed to improve access to screening and diagnosis in underserved regions.

MindBricks / Solvio

React · API integration · Language learning

Contributed to an AI-powered language-learning product integrating frontend workflows with the MindBricks microservice ecosystem.

EXPERIENCE & ORGANIZATIONS

Ethosoft · Co-Founder · Strategy Lead · AI & Data Engineer

2024–Present

Istanbul, Türkiye · Student-led AI research and software organization

- Co-founded and helped develop Ethosoft from the earlier Ethos initiative into a public-facing research and software team.
- Works across Turkish NLP, small language models, document AI, medical AI, AI safety, robotics, space systems, and applied AI products.
- Helped establish a flat, ownership-driven working model: concepts are developed collaboratively, while one project lead owns execution, testing, and delivery.
- Published public preprints, models, datasets, GitHub repositories, and technical reports through personal and Ethosoft accounts.
- Used TRUBA national research-computing resources for eligible projects and supported research reproducibility through scripts, manifests, result tables, and documentation.
- Presented team work at Mersin University R&D Project Market 2026 and contributed to a research culture connecting high school builders with academic and open-source communities.

VolTrAN Robotics #71080V · Software Captain · Software & Mechanical Team Member

2023–Summer 2026

Pertevniyal High School · VEX robotics

- Progressed from trainee/member to software leadership while contributing across robot programming, autonomous systems, CAD, technical drawing, assembly, scouting tools, and mechanical prototyping.
- Led a four-member software department and coordinated approximately 15 junior/subteam members through member onboarding, software planning, autonomous development, and competition preparation.
- Delivered weekly online and in-person instruction, with intensive after-school support, covering C++, reusable/template code, PID control, path planning, and autonomous robot programming.
- Helped keep the team active and rebuild its technical and organizational capacity when it faced possible closure during the 2023–24 and 2024–25 seasons.
- Contributed to the foundation for a later 5–1 qualification run and 7th-place ranking at the 2026 Türkiye national event.
- Worked during a season in which VolTrAN recorded the tournament's highest autonomous-software score at the 2025 International VEX Open in Pendik.
- Co-organized the Circle of Awareness social-responsibility initiative, conducting outreach meetings with approximately 15 VEX teams.

Uyumsoft · Contract AI / Data Science Contributor

August 2025

Istanbul, Türkiye · Paid one-month contractual project · Two-person team

- Worked on an enterprise AI-integration project to forecast the following month's sales from the preceding 12 months of historical sales data.
- Owned model training and contributed to transforming multidimensional business data into a forecasting-ready time series.
- Trained and delivered an XGBoost forecasting model and completed the project at the end of the one-month engagement.

Neuronauts AI Team · Co-Founder

2024–2025

Pertevniyal High School

- Co-founded a student AI team focused on machine learning, deep learning, neural networks, technical talks, and project-based learning.
- Co-organized the Pertevniyal Artificial Intelligence Summit, featuring speakers from Google on computer vision and Trendyol on natural language processing.
- Supported volunteer Python and AI education for incoming student cohorts.
- Recognized organizational sustainability risks and helped establish Ethos, which later developed into Ethosoft.

Kosi Metal · Manufacturing & Technical Drawing Experience

Summer 2024

Istanbul, Türkiye

- Gained hands-on exposure to CNC laser cutting, sheet-metal bending, welding, assembly, technical drawing, and production constraints.
- Applied robotics-based CAD and mechanical design skills in an industrial setting and later used this experience in medical-device prototyping.

Bilişim Vadisi GO Istanbul · Entrepreneurship Ecosystem Participant

2026–Present

- Continued Ethosoft research and product development within an entrepreneurship ecosystem focused on moving technical prototypes toward real-world use.

LEADERSHIP & COMMUNITY ENGAGEMENT

Technical Education & Mentoring Organized recurring online, in-person, and after-school training in C++, PID control, path planning, code templates, autonomous robotics, Python, and introductory AI for younger students.	Circle of Awareness Co-organized a VEX community social-responsibility initiative and participated in outreach discussions with approximately 15 robotics teams.
Pertevniyal AI Summit Co-organized a school AI summit featuring industry speakers on computer vision and natural language processing.	Academic & Research Outreach Connected student research with ITU lectures, TRUBA computing, public research releases, Mersin University R&D Project Market, and RoboCup-related team activity.

SELECTED AWARDS & COMPETITIONS

11 Mar 2026	EU Code Week Hackathon Finals — Jury Award · Head of Mechanics Led the mechanical and prototyping side of Ethosoft's portable AI-driven eye-care platform, applying CAD, assembly, robotics, and manufacturing experience to a medical AI system designed for underserved communities.
15–17 May 2026	DataMedX Hackathon 2 — 2nd Place Developed an oncology clinical AI orchestration and doctor-panel system with patient retrieval, specialized analysis agents, evidence-backed summaries, risk cards, timelines, and report generation.
Feb 2026	MindBricks Hackathon 3 — 1st Place · Ethosoft Contributed to Ethosoft's product and technical development in a competition centered on building AI-powered applications with the MindBricks ecosystem.
Mar 2026	TUA Astro Hackathon — Istanbul European Side Regional Champion Built a lunar-rover navigation project combining slope-constrained 3D A* global planning, SAC local control, simulated sensing, and NASA lunar topography.
10–11 Apr 2026	Harvard Health Systems Innovation Lab Hackathon · Istanbul Hub Finalist Developed BioDockX, an AI-agent bioinformatics platform for protein retrieval, interactive 3D visualization, ligand preparation, and molecular docking automation.
23–25 May 2025	MostrAI Hackathon — Special Jury Award · Sirius Team Competed in a 48-hour AI/software hackathon; the result led to a formal paid contractual AI project opportunity at Uyumsoft.
2024 & 2026	Türkiye VRC Finals — Quarterfinalist Seasons · VolTrAN Robotics Contributed to software, autonomous programming, scouting, CAD, assembly, and team preparation across national VEX competition seasons.
2026	VRC European Open — Inspire Award Season · VolTrAN Robotics Participated in the team's software and competition preparation during its European Open season.

TECHNICAL SKILLS

Programming & Systems

Python C / C++ JavaScript TypeScript Bash Linux
Git / GitHub Docker SQLite

Machine Learning & LLMs

PyTorch Hugging Face Transformers QLoRA / PEFT
bitsandbytes SFT XGBoost Triton GGUF

NLP & Research Methods

Autoregressive LMs Tokenizer diagnostics Semantic reranking
Domain transfer Bootstrap CIs Multiple-choice evaluation
Reproducibility

HPC & Infrastructure

TRUBA SLURM H200 GPU jobs Ollama Model serving
Compute accounting Checkpoint validation

Robotics & Mechanics

VEX V5 Autonomous programming PID control Path planning
3D A* SAC CAD Technical drawing Mechanical assembly

Web, Medical AI & Bioinformatics

FastAPI React / Vite REST APIs Clinical AI agents RCSB PDB
PubChem NGL Viewer Open Babel AutoDock Vina

RESEARCH INFRASTRUCTURE & REPRODUCIBILITY

TRUBA research computing SLURM jobs H200 GPU workflows Reproducibility manifests Checkpoint hashing
Bootstrap confidence intervals Compute-cost accounting Model and dataset cards Technical-report writing

RESEARCH COMMUNICATION

Writing

Research manuscripts, technical reports, model and dataset cards, READMEs, experiment registries, and reproducibility documentation.

Presentation

Hackathon pitching, project-market presentations, technical demonstrations, research outreach, and robotics judge presentations.

Teaching

Recurring instruction in programming, autonomous robotics, PID, path planning, Python, and introductory AI.

LANGUAGES & PROFESSIONAL LINKS

Languages

- Turkish — Native
- English — English preparatory education; EnglishScore 458

Selected Profiles

- github.com/Ahmet2001
- huggingface.co/Ethosoft
- orcid.org/0009-0001-2941-0412
- ethosoft.org