# Bechir Dardouri

bechir.dardouri@ept.ucar.tn | bechir-dardouri | Bechirdardouri | Scholar

## Profile

What keeps me in this field is the stranger question of how a model learns to think efficiently. The decisions hidden inside that question are what I find most interesting: what we teach, how we teach it, what we let it be wrong about. I work mostly where reasoning has to be checked rather than taken on faith, and I move between research and engineering with a stubborn belief that the next interesting wave of intelligence will come from how we shape models.

## Education

**University of Tübingen — Tübingen AI Center** Tübingen, Germany
*M.Sc., Applied Mathematics & Machine Learning — Research advisor: Prof. Hildegard Kühne* *Apr. 2026 – Present*
- Working at the Tübingen AI Center, one of Europe's leading AI hubs, with research focused on efficient multimodal foundation models, vision-language compression, and post-training methods for omni-modal architectures

**École Polytechnique de Tunisie** La Marsa, Tunisia
*Engineering Degree, Applied Mathematics & Computer Science* *Sep. 2024 – Jun. 2027 (expected)*
- Selective grande école in the French engineering tradition, blending rigorous applied mathematics with computer science. Coursework: Advanced Deep Learning, Convex & Stochastic Optimization, Probability and Statistics, Distributed Systems, Numerical PDEs and Variational Methods, Information Theory

**Institut Préparatoire aux Études d'Ingénieur de Monastir** Monastir, Tunisia
*Classes Préparatoires — Mathematics, Physics & Computer Science (MP option)* *Sep. 2022 – Jun. 2024*
- Two-year intensive scientific program (Mathematics, Physics, Theoretical Computer Science, Engineering Sciences) preparing for the national *concours* to the Tunisian engineering grandes écoles. National ranking: top 1.2%

## Experience

**Tübingen AI Center** Tübingen, Germany
*Visiting Researcher — Kühne Group* *Mar. 2026 – Present*
- Collaborating with Sofiane Chaybouti on token compression for vision-language and omni-modal foundation models—designing architectures that scale perceptual context without scaling sequence length, and studying how post-training recipes keep compressed representations faithful to fine-grained reasoning

**Tanit Healthcare Technologies** Paris, France
*Research Engineer — Medical Foundation Models* *Jun. 2025 – Present*
- Lead post-training of in-house medical Large Reasoning Models with a recipe built around knowledge distillation from a curated in-house medical knowledge graph, curriculum learning, and RL with verifiable rewards over multi-hop reasoning paths—developed in collaboration with leading French research institutions to advance the state of the art in medical LLMs

**Hamad bin Khalifa University (HBKU)** Doha, Qatar
*Research Engineer Intern — QCRI Social Innovation Program* *Jun. 2025 – Oct. 2025*
- Designed RoBee, a multi-agent neuro-symbolic pipeline that turns raw qualitative interviews into validated causal graphs through SCM-guided prompts, counterfactual stress-tests, and structural pruning—producing interpretable structures for policy analysis and downstream decision support

## Publications

**VulnScout-C: A Lightweight Transformer for C Code Vulnerability Detection** | arXiv:2603.28309 2026
- A 693M-parameter Mixture-of-Experts transformer (353M active) derived from Qwen3, paired with a 33K-sample dual-verified dataset built through a multi-agent pipeline. Outperforms GPT-4o, GPT-o3 Mini, and DeepSeek R1 on the CASTLE benchmark at a fraction of their inference cost—under submission to IEEE TDSC

**TTFT-Aware Graph Chain-of-Thought** | *MEDES 2025 (Springer, Jan. 2026)* | Paper 2025
- Distance-indexed Neural A* for multi-hop medical reasoning—auditable, low-latency clinical decision paths

**CSG-Chat: Constructing Causal Structured Graphs** | *QCRI-SIP 2025* | Poster 2025
- Multi-agent neuro-symbolic pipeline for automated causal graph construction from qualitative interviews

## Selected Projects

**GOAT_AI 1.0 — Monocular Depth Estimation, 2nd place** | Repository 2026
- Reverse-engineered the competition's ground truth as a Depth Anything V2-Large output and deliberately chose the smaller V2-Small as the distillation teacher—turning a redundant-signal problem into a complementary one, in a 1M-parameter, 4 MB student

**Papers QA** | Repository 2025
- RAG system for research papers with vector embeddings and semantic search over LangChain—fast, contextual answers across long technical PDFs