

Spyros Mouselinos | Applied ML Researcher & Technical Leader

Amsterdam, Netherlands

M +31 6 15061443 • E mouselinos.spur.kw@gmail.com

W spyrosmouselinos.github.io/CV/ • gh SpyrosMouselinos • Google Scholar

Applied machine-learning researcher and technical leader specializing in multimodal generative AI, agentic systems, and production ML. Built zero-to-one multi-agent systems for eBay Ads; previously developed frontier text-to-video models at Moonvalley and led production ML systems serving 20M predictions per day. First author at ICLR and Findings of ACL and EMNLP, with end-to-end ownership across research, data, evaluation, distributed inference, and deployment.

Experience

eBay

Amsterdam, Netherlands

Senior Applied Researcher, Ads

2026–Present

- Architected and built multiple multi-agent systems from the ground up for eBay Ads, spanning recommendation, advertising-opportunity discovery, and automation of internal technologies.
- Owned zero-to-one development from problem formulation and agent-role design through orchestration and implementation, turning ambiguous research ideas into working systems rather than configuring an off-the-shelf agent workflow.
- Applied agentic approaches across distinct Ads use cases to surface recommendations and automate multi-stage internal technical workflows.

Moonvalley

London, Remote

Member of Technical Staff, Foundational AI

2024–2026

- Led R&D initiatives improving rendered-text fidelity and temporal consistency in text-to-image and text-to-video foundation models, including motion-aware placement and style transfer on moving surfaces.
- Developed preference-optimization workflows for 70B-parameter LLM/VLM assistants, aligning model outputs with creative intent and reducing downstream iteration for studio workflows.
- Designed synthetic-data labeling and importance-sampling pipelines for hundreds of millions of licensed video-text pairs, expanding coverage of rare concepts and diverse motion regimes under strict data-provenance constraints.

Equinox AI B.V.

The Hague, Netherlands

Consulting ML Architect

Jun 2024–Sep 2024

- Set the architecture for a low-latency algorithmic-trading platform, integrating real-time order-book and alternative-data feeds with online inference, Redis caching, statistical risk controls, and automated kill switches.
- Built Bayesian-updated reinforcement-learning agents and a distributed Monte Carlo backtesting framework for hot-swap rollouts; introduced design reviews, runbooks, and Grafana observability before handing the platform to the internal team.

DeepSea Technologies

Athens, Greece

ML Engineer / Technical Lead

2018–2021

- Served as technical lead for a five-engineer ML platform team; established containerized deployment, experiment tracking, CI/CD, blue-green releases, monitoring, and over-the-air model updates.
- Productionized a multimodal Transformer serving 20M predictions per day and automated anomaly detection that reduced bunker-fuel fraud by 50% across a fleet of 1,000 vessels.
- Built a deep-learning-assisted graph route optimizer used for real-time routing of 100+ vessels, outperforming contemporary competitor solutions by 8–10%.
- Developed vessel-performance and predictive-maintenance models that reduced fuel consumption by 4% and maintenance costs by 10% across customer deployments.

Selected Venture Experience

Synthesa AI

Remote

Founding ML Architect

2025–Present

- Designed the core cloud-native LLM platform for pharmaceutical and biotech evidence synthesis, combining Django, Docker, GCP, distributed queues, KV caching, and vLLM serving for predictable, traceable execution.
- Built parallel agentic workflows that coordinate language and vision-language models to ingest, retrieve, and reason over scientific documents; implemented multimodal RAG with vector search, reranking, and policy filters.
- Contributed the ML architecture behind a platform whose published validation screened 270,626 abstracts with 100% sensitivity and 99.4% specificity while reducing manual-review workload by 91.7%.

Research & Academic Leadership

University of Warsaw

Warsaw, Poland

Doctoral Researcher & Teaching Assistant

2021–2026

- Established a research program on reasoning and robustness in language and vision-language models, resulting in first-author publications at ICLR, Findings of ACL, and Findings of EMNLP.
- Developed black-box adversarial evaluation, bias-discovery, and multi-agent reasoning methods that exposed failure modes in visual reasoning, code generation, and constructive geometry systems.
- Taught master's-level NLP and Visual Recognition laboratories and mentored student teams on research-oriented vision-language projects.

Selected Publications

EMNLP 2024: S. Mouselinos, H. Michalewski, M. Malinowski. "Beyond Lines and Circles: Unveiling the Geometric Reasoning Gap in Large Language Models." Findings of EMNLP.

ACL 2023: S. Mouselinos, M. Malinowski, H. Michalewski. "A Simple, Yet Effective Approach to Finding Biases in Code Generation." Findings of ACL.

ICLR 2022: S. Mouselinos, H. Michalewski, M. Malinowski. "Measuring CLEVRness: Black-box Testing of Visual Reasoning Models."

IJCNN 2021: "MAIN: Multihead-Attention Imputation Networks."

Technologies 2020: "Optimized Generation of Hardware CNN Inference Engines."

IEEE MOCAST 2019: "TF2FPGA: A Framework for Projecting and Accelerating CNNs on FPGA Platforms."

Education

University of Warsaw

Poland

PhD, Computer Science

2026

Dissertation: *Visual and Language Reasoning in Deep Learning Models*

Athens University of Economics and Business

Greece

MSc, Data Science

2021

Focus: language models, conversational AI, and data mining

National Technical University of Athens

Greece

BSc & MEng, Electrical and Computer Engineering

2018

Focus: software systems, computer architecture, and artificial intelligence

Technical Expertise

Research: Multimodal generative AI; agentic and multi-agent systems; LLM/VLM alignment; recommendation; retrieval and RAG; reinforcement learning; evaluation and robustness; synthetic data

ML systems: PyTorch; TensorFlow; Hugging Face Transformers; vLLM; Ray/RLlib/Serve; ONNX; MLflow; Weights & Biases; Spark; Dask

Platform: Python; SQL; Docker; Kubernetes; GCP; AWS; FastAPI; Jenkins; Grafana; PostgreSQL; Redis; MongoDB; InfluxDB; FAISS

Languages

Greek (native) English (C2) German (B1)