

GERASIMOS CHATZOUDIS

✉gerasimos.chatzoudis@rutgers.edu | 🌐yerassimos.com | 🌐yerassimos | 🌐gerasimos-chatzoudis

Education

Rutgers University – The State University of New Jersey Piscataway, NJ
MS/ PhD in Computer Science: Mechanistic Interpretability on Vision-Language Models GPA: 4.0/4.0
Advisor: Prof. Dimitris N. Metaxas Expected: August 2026
National Technical University of Athens Athens, Greece
Joint MEng & BSc, Electrical and Computer Engineering — Major in Computer Science July 2020

Publications

1. **Chatzoudis, G.**, Polyzos, K.D., Li, Z., Gu, D., Moran, G.E., Wang, H., Metaxas, D.N. (2026). *Can Cross-Layer Transcoders Replace Vision Transformer Activations? An Interpretable Perspective on Vision*. **XAI4CV Workshop, CVPR 2026 (Spotlight)**. [project page]
2. **Chatzoudis, G.**, Li, Z., Moran, G.E., Wang, H., Metaxas, D.N. (2025). *Beyond Interpretability: When, Why, and How Sparse Autoencoders Enable Label-Free Visual Steering*. arXiv preprint.
3. **Chatzoudis, G.**, Plitsis, M., Stamouli, S., Dimou, A.-L., Katsamanis, A., Katsouros, V. (2022). *Zero-Shot Cross-Lingual Aphasia Detection using Automatic Speech Recognition*. **INTERSPEECH 2022**.
4. Gu, D., Gao, Y., **Chatzoudis, G.**, Dong, Z., Zhang, G., Guo, B., Zhou, Y., Zhou, M., Metaxas, D.N. (2025). *LUCID-SAE: Learning Unified Vision-Language Sparse Codes for Interpretable Concept Discovery*. arXiv preprint.
5. Stamouli, S., . . . , **Chatzoudis, G.**, et al. (2023). *A Web-Based Application for Eliciting Narrative Discourse from Greek-Speaking People with Language Impairments*. **Frontiers in Communication**.

Research Experience

Rutgers University – Computational Biomedicine Imaging and Modeling Center Piscataway, NJ
PhD Candidate, advised by Prof. Dimitris N. Metaxas Sep 2022 – Present

- Developed **Cross-Layer Transcoders (CLTs) for mechanistic interpretability in Vision Transformers**, enabling layer-wise attribution of model predictions to sparse, human-interpretable visual circuits across ViT-B/16 and ViT-B/32; accepted at XAI4CV Workshop, CVPR 2026.
- Designed **Visual Sparse Steering (VS2)**, the first systematic investigation of unsupervised SAE sparse features for downstream vision steering; *improves CLIP zero-shot classification* across 9 datasets and 2 backbones with per-class gains up to +25% (VS2) and +38% (VS2+); *interpretable by design, efficient* (<0.1% compute overhead vs. ~16.6× for prompt-based TTA, 1,800× cheaper than LoRA), and *reliable* via a FVU-gated fallback safety mechanism absent from standard steering methods; arXiv preprint
- Collaborated on **LUCID-SAE**, developing *unified vision-language sparse codes* for interpretable concept discovery across modalities; arXiv 2025.

Athena Research Center Athens, Greece
Machine Learning Scientist Sep 2020 – June 2022

- Built a cross-lingual aphasia detection system achieving **88% zero-shot accuracy** on low-resource languages via transfer learning from high-resource speech models; published at INTERSPEECH 2022.
- Built and deployed a React/FastAPI web application for remote and in-person narrative discourse elicitation from patients with aphasia, featuring real-time audio streaming via RabbitMQ; deployed across clinical settings with speech-language pathologists
- Managed a 25-member Speech-Music-Signal-Dialogue team as Scrum Master, coordinating cross-functional research delivery

Speech and Language Processing group at NTUA Athens, Greece
Research Assistant Sep 2019 – June 2020

- Master's thesis (10/10): developed a modular negotiation dialogue system (BERT + GPT-2) on the Deal or No Deal task, applying Reinforcement Learning to decouple agent strategies and demonstrating convergence toward social equilibrium; advised by Prof. Alexandros Potamianos and Dr. Athanasios Katsamanis.

Industry Experience

Amazon San Diego, California
Applied Scientist Intern (L5) May 2025 – Aug 2025

- Built a scalable knowledge graph construction pipeline over large-scale behavioral data, developing novel multi-pass triplet extraction methods improving F1 by ~25% over single-pass baselines
- Created an annotated behavioral-to-KG dataset with a custom labeling UI, establishing ground truth for pipeline evaluation
- Developed an LLM-based code generation system for triplet extraction, replacing direct LLM inference and compressing raw data into structured representations, yielding substantial estimated infrastructure cost savings
- Engineered a dual GraphRAG retrieval system combining graph traversal and semantic similarity indexing, with an agentic LLM router to select the optimal retrieval strategy per query
- Demonstrated production viability across targeting, fraud, and segmentation use cases

Amazon Web Services (AWS) Santa Clara, California
Applied Scientist Intern (L5) May 2024 – Aug 2024

- Designed and implemented GraphRAG methods for multi-hop reasoning, achieving measurable gains over standard dense retrieval baselines on complex multi-step queries

Innoetics Text-To-Speech Technologies (Samsung Greece) Athens, Greece
Machine Learning Intern Apr 2019 – June 2019

- Implemented emotional expressivity control in a Tacotron-based TTS system for more natural speech synthesis

Skills

Research Areas	Mechanistic Interpretability, Sparse Autoencoders, Vision Transformers, CLIP, LLMs, GraphRAG, Knowledge Graphs, Transfer Learning, Unsupervised Adaptation, Steering Vectors, LLM Reasoning
Programming Languages	Python, C/C++, SQL, Bash, MATLAB, JavaScript, React, HTML, CSS
Frameworks & Libraries	PyTorch, TensorFlow, Hugging Face, LangChain, FastAPI, Flask, Streamlit, React, Scikit-learn, Pandas, NumPy, OpenCV
Experiment Tracking & Optimization	Weights & Biases, MLFlow, Optuna, TensorBoard
DevOps & Cloud	Docker, Git, RabbitMQ, AWS, GCP
Languages	Greek (native), English (proficient), German (C1)

Teaching

Teaching Assistant, Rutgers University: Introduction to AI (Fall 2022, Spring 2025) · Design & Analysis of Algorithms (Spring & Fall 2024) · Database Systems for Data Science (Spring 2023) · Topics in Computers in Biomedicine (Fall 2023)

Leadership and Community Involvement

TEDxNTUA · Co-Curator, Head of Speakers Team, Volunteer · Athens, Greece · 2017–2019

- Curated 23 speakers across two events; led a 28-person organizing team

Honors and Awards

- Gerondelis Foundation Scholarship (2023)
- Stavros Niarchos Foundation Scholarship (2019–2020)
- Deree – The American College of Greece Scholarship (2019–2020)
- Top 1% nationwide, Greek National University Admissions Exams (2014)
- Excellence Award, Mayor of Municipality of Pallini (2014)
- 9th place, World Youth Chess Championship (2006); two-time A-Division competitor; 3rd Mediterranean U20 Chess Championship; World Junior U20 Chess Championship