

Georgios Pantazopoulos

✉ gmp2000@hw.ac.uk • 🌐 <https://gpantaz.github.io/> • 🐙 github • 🎓 scholar
in linkedin • 🐦 X

Education

PhD in Robotics and Autonomous Systems

Sep 2021 - present

CDT in Robotics and Autonomous Systems, Heriot-Watt University, University of Edinburgh

- Thesis: Inductive Biases in Architectures of Modern Visual Language Models, supervised by [A. Suglia](#)

MSc in Artificial Intelligence with Speech and Multimodal Interaction

Sep 2020 - Aug 2021

Heriot-Watt University

- Grade: 84/100 (1st with Distinction)
- Thesis: Compositional Generalization in Image Captioning, supervised by [A. Eshghi](#)

MEng in Electrical and Computer Engineering

Sep 2012 - Jun 2018

National Technical University of Athens

- Grade: 7.7/10, Major: Computer Science
- Thesis: Stress Detection from Speech Signals, supervised by [A. Potamianos](#)

Publications

Evaluating Multimodal Language Models as Visual Assistants for Visually Impaired Users

2025 Annual Meeting of the Association for Computational Linguistics, [\[paper\]](#)

- A.Karamolegkou*, M.Nikandrou*, **G.Pantazopoulos***, D.Sanchez, P.Rust, R.Dhar, D.Hershcovich, A.Søgaard

CROPE: Evaluating In-Context Adaptation of Vision and Language Models to Culture-Specific Concepts

2025 Annual Conference of the North American Chapter of the Association for Computational Linguistics, [\[paper\]](#)

- M. Nikandrou, **G. Pantazopoulos**, N. Vitsakis, I. Konstas, A. Suglia

Shaking Up VLMs: Comparing Transformers and Structured State Space Models for Vision & Language Modeling

2024 Conference on Empirical Methods in Natural Language Processing, [\[paper\]](#)

- G. Pantazopoulos**, M.Nikandrou, A. Suglia, O. Lemon, A. Eshghi

Enhancing Continual Learning in Visual Question Answering with Modality-Aware Feature Distillation

2024 ACL Third Workshop on Advances in Language and Vision Research, [\[paper\]](#)

- M.Nikandrou*, **G. Pantazopoulos***, I. Konstas, A. Suglia

Lost in Space: Probing Fine-grained Spatial Understanding in Vision and Language Resamplers

2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics, [\[paper\]](#)

- G. Pantazopoulos**, A. Suglia, O. Lemon, A. Eshghi

Learning To See But Forgetting To Follow: Visual Instruction Tuning Makes LLMs More Prone To Jailbreak Attacks

2024 LREC-COLING Third Workshop on Safety for Conversational AI, [\[paper\]](#)

- G. Pantazopoulos**, A. Parekh, Malvina Nikandrou, A. Suglia

Multitask Multimodal Prompted Training for Interactive Embodied Task Completion

2023 Conference on Empirical Methods in Natural Language Processing, [\[paper\]](#)

- G. Pantazopoulos**, M.Nikandrou, A. Parekh, B. Hemanthage, A. Eshghi, I. Konstas, V. Rieser, O. Lemon, A. Suglia

EMMA: A foundation model for embodied, interactive, multimodal task completion in 3D environments

2023 Alexa Prize SimBot Challenge Proceedings, [\[paper\]](#)

- A. Parekh*, M.Nikandrou*, **G. Pantazopoulos***, B. Hemanthage, A. Eshghi, I. Konstas, O. Lemon, A. Suglia

Combine to Describe: Evaluating Compositional Generalization in Image Captioning

2022 ACL Student Research Workshop, [\[paper\]](#)

- G. Pantazopoulos**, A. Suglia, A. Eshghi

Data Augmentation using GANs for Speech Emotion Recognition

2019 Annual Conference of the International Speech Communication Association, [\[paper\]](#)

- A. Chatziagapi, G. Paraskevopoulos, D. Sgouropoulos, **G. Pantazopoulos**, M.Nikandrou, T. Giannakopoulos, A. Katsamanis, A. Potamianos, S. Narayanan

Projects

Retrieval

Jan 2025 - Jun 2025

Heriot-Watt University

- Benchmarking Transformers, State Space Models, and Hybrid models on long sequence modeling

- EMMA** Jan 2022 - May 2023
Alexa Prize Simbot Challenge, Finalist [code]
- Developed an interactive language-guided agent executing instructions in simulated environments
- Exploring the Role of Patch Resolution for Vision Transformers** Jan 2022 - Mar 2022
University of Edinburgh
- Explored the trade-off between performance gains and inference cost on downstream tasks

Industry Experience

- AI Research Intern** Jun 2025 - present
Alan Turing Institute
- Worked on thinking models for GUI agents
 - Developed ranking models for filtering noisy data for training GUI agents
- Software Engineer Intern** Feb 2025 - Apr 2025
Reka AI
- Implemented constrastive methods for image-to-image, image-to-video, image & text-to video retrieval
- Machine Learning Engineer** Sep 2018 - Apr 2020
Behavioral Signal Technologies Inc.
- Developed audio and text models for emotion recognition
 - Established pipelines for data collection and annotation from crowdsourcing platforms
 - Contributed to the implementation of voice activity detection algorithms
 - Carried out research in active learning pipeline to address class imbalances
 - Contributed to real-time systems for emotion and hesitation recognition

Teaching Experience

- AθNLP 2024 / 2025** Sep 2024 / Sep 2025
Summer School Teaching Assistant, [code]
- Developed the lab materials for sequence modeling and multimodal models
- Data Mining and Machine Learning** Sep 2022 - present
Teaching Assistant
- Contributed to the infrastructure and the coursework of the module
 - Supervised groups of postgraduate students on their semester project
- Introduction to Natural Language Processing** Sep 2023 - present
Teaching Assistant, [code]
- Contributed to the infrastructure and the coursework of the module
 - Supervised groups of students on their semester project

Technical skills

- | | |
|------------------------------|------------------------------------------------------------------------------------|
| Programming Languages | Python, Matlab, R, C, C++, C#, Java, PDDL, Bash, SML |
| Frameworks | PyTorch, Hugging Face, PyTorch Lightning, Faiss, Milvus, Trax, Keras, scikit-learn |
| Development Tools | Weights & Biases, Gradio, AWS, Git, Slurm, Docker, Kubernetes, Jenkins |

Invited Talks

- An Efficient Training Pipeline for Reasoning Graphical User Interface Agents** Nov 2025
Alan Turing Institute, [slides]
- Embodied MultiModal Agent for Language-guided Action Execution in 3D Simulated Environments** Oct 2022
CDT in Robotics and Autonomous System Conference, [slides]

Honors & Awards

- Evaluating Multimodal Language Models as Visual Assistants for Visually Impaired Users** Jul 2025
Pioneer Center of AI Best Paper Award, ACL 2025 SAC Highlights
- Exploring the Role of Patch Resolution for Vision Transformers** Mar 2022
University of Edinburgh, 2nd place in the annual IBM Awards

Academic Service

- ARR Reviewer** Jun 2024 / Oct 2024 / Jun 2025
- IEEE SLT218 Conference** Dec 2018
Conference Volunteer