

# Georgios Pantazopoulos

✉ gmp2000@hw.ac.uk • 🌐 george-pantazopoulos • 📧 george-pantazopoulos  
in george-pantazopoulos • 🐦 george-pantazopoulos

## Education

### PhD in Robotics and Autonomous Systems

Sep 2021 - present

*Centre for Doctoral Training in Robotics and Autonomous Systems, Heriot-Watt University*

- 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: Combine to Describe: Evaluating Compositional Generalisation in Image Captioning, supervised by [Arash 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

### ViCA: Combining visual, social, and task-oriented conversational AI in a healthcare setting

2021 23rd ACM International Conference on Multimodal Interaction, [\[paper\]](#)

- **G. Pantazopoulos**, J. Bruyere, M.Nikandrou, Thibaud Boissier, S. Hemanthage, S. Binha, V. Shah, C. Dondrup, O. Lemon

### 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

|   |                      |
|---|----------------------|
| <b>Retrievit</b><br><i>Heriot-Watt University</i>   | Jan 2025 - June 2025 |
| <ul style="list-style-type: none"><li>◦ Benchmarking Transformers, State Space Models, and Hybrid Architectures on long sequence modeling using synthetic tasks</li></ul> |                      |
| <b>EMMA</b><br><i>Alexa Prize Simbot Challenge, Finalist <a href="#">[code]</a></i>   | Jan 2022 - May 2023  |
| <ul style="list-style-type: none"><li>◦ Developed an interactive language-guided agent executing instructions in simulated environments</li></ul>                         |                      |
| <b>Exploring the Role of Patch Resolution for Vision Transformers</b><br><i>University of Edinburgh, 2nd place in the annual IBM Awards</i>                               | Jan 2022 - Mar 2022  |
| <ul style="list-style-type: none"><li>◦ Explored the trade-off between performance gains and inference cost on downstream tasks</li></ul>                                 |                      |

## Industry Experience

|   |                       |
|---|-----------------------|
| <b>AI Research Intern</b><br><i>Alan Turing Institute</i>   | June 2025 - present   |
| <ul style="list-style-type: none"><li>◦ Working on thinking models for image-level understanding</li></ul>  |                       |
| <b>Software Engineer Intern</b><br><i>Reka AI.</i>  | Feb 2025 - April 2025 |
| <ul style="list-style-type: none"><li>◦ Worked on image-to-image, image-to-video, image &amp; text-to video retrieval</li></ul>   |                       |
| <b>Machine Learning Engineer</b><br><i>Behavioral Signal Technologies Inc.</i>  | Sep 2018 - Apr 2020   |
| <ul style="list-style-type: none"><li>◦ Developed audio and text models for emotion recognition</li><li>◦ Established pipelines for data collection and annotation from crowdsourcing platforms</li><li>◦ Contributed to the implementation of voice activity detection algorithms</li><li>◦ Carried out research in active learning pipeline to address class imbalances</li><li>◦ Contributed to real-time systems for emotion and hesitation recognition</li></ul> |                       |

## Teaching Experience

|  |                     |
|--|---------------------|
| <b>A@NLP 2024 / 2025</b><br><i>Summer School Teaching Assistant, <a href="#">[code]</a></i>  | Sep 2024 / Sep 2025 |
| <ul style="list-style-type: none"><li>◦ Developed the lab materials for sequence modeling and multimodal models</li></ul>  |                     |
| <b>Data Mining and Machine Learning</b><br><i>Teaching Assistant</i>   | Sep 2022 - present  |
| <ul style="list-style-type: none"><li>◦ Supervised groups of postgraduate students on their semester project</li></ul>   |                     |
| <b>Introduction to Natural Language Processing</b><br><i>Teaching Assistant</i>  | Sep 2023 - present  |
| <ul style="list-style-type: none"><li>◦ Contributed to the infrastructure and the coursework of the module</li><li>◦ Supervised groups of students on their semester project</li></ul> |                     |

## Technical skills

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

## Invited Talks

|  |          |
|--|----------|
| <b>CDT in Robotics and Autonomous Systems 2022 Conference</b><br><i>Heriot-Watt University / University of Edinburgh, <a href="#">[slides]</a></i> | Oct 2022 |
|--|----------|

## Academic Service

|  |                     |
|--|---------------------|
| <b>ARR Reviewer</b>  | Jun 2024 / Oct 2024 |
| <b>IEEE SLT218 Conference</b><br><i>Conference Volunteer</i> | Dec 2018            |