

# Tobia Poppi

PHD CANDIDATE · RESPONSIBLE AI RESEARCHER

Modena, Italy

✉ tobia.poppi@gmail.com | 🏠 tobiapoppi.github.io | 📷 tobiapoppi | 📺 tobia-p-890793140

“Empowering AI to be a Safe, Trustworthy, and Robust tool for Society.”

PhD Candidate in Artificial Intelligence specializing in **Responsible AI**, **Generative Models**, and **Multimodal Learning**. Recipient of the **Amazon Research Fellowship**. Author of **CVPR (Highlight)** and **ECCV** publications. Currently working on safety-alignment pipelines and synthetic data workflows to mitigate hallucinations and harmful content in Vision-Language Models (VLMs). Driven by a commitment to engineering AI systems that prioritize safety, reliability, and alignment with human values in real-world applications.

## Education

### University of Pisa (National PhD in AI for Society)

Modena, Italy

PHD IN ARTIFICIAL INTELLIGENCE

Nov. 2023 - Nov. 2026 (expected)

- Focus: **Responsible Generative AI**, specializing in safety alignment, robustness, and transparency for Multimodal Systems.
- Host Institution: AlmageLab, University of Modena and Reggio Emilia.

### University of Modena and Reggio Emilia

Modena, Italy

MASTER'S DEGREE IN COMPUTER ENGINEERING

Sep. 2021 - Oct. 2023

- **Grade: 110/110 cum laude**
- Erasmus+ Exchange: Norwegian University of Science and Technology (NTNU), Trondheim (Aug. 2022 - Dec. 2022).

### University of Modena and Reggio Emilia

Mantova, Italy

BACHELOR'S DEGREE IN COMPUTER ENGINEERING

Sep. 2018 - Oct. 2021

- Grade: 105/110

## Selected Publications

### Hyperbolic Safety-Aware Vision-Language Models (CVPR 2025)

Nashville, USA

TOBIA POPPI, TEJASWI KASARLA, PASCAL METTES, LORENZO BARALDI, RITA CUCCHIARA

2025

- **CVPR 2025 Highlight** (Top ~2.8% of submissions).
- Introduced **HySAC**, a hierarchical safety-aware CLIP model utilizing hyperbolic embeddings to dynamically redirect unsafe content retrieval to safer alternatives, achieving SOTA on safety benchmarks without compromising utility.

### Safe-CLIP: Removing NSFW Concepts from Vision-and-Language Models (ECCV 2024)

Milan, Italy

SAMUELE POPPI, TOBIA POPPI, FEDERICO COCCHI, MARCELLA CORNIA, LORENZO BARALDI, RITA CUCCHIARA

2024

- Developed a fine-tuning strategy to dissociate NSFW concepts in VLMs, ensuring safe deployment for text-to-image retrieval and generation.

### Counterfactual Video Generation for Mitigating Action and Temporal Hallucinations in Video-Language Models

TOBIA POPPI, BURAK UZKENT, AMANMEET GARG, LUCAS PORTO, GARIN KESSLER, YEZHOU YANG, MARCELLA CORNIA, LORENZO

Under Review (2026)

BARALDI, RITA CUCCHIARA, FLORIAN SCHIFFERS

- Proposed a pipeline using agentic VLM reasoning and image/video diffusion to generate counterfactual training pairs, targeting specific VLM hallucinations on video (objects, actions, temporal).
- Work conducted at Amazon Prime Video; publication approved via Amazon SIM process.

## Experience

### Amazon Prime Video

Seattle, USA

APPLIED SCIENTIST INTERN

Jul. 2025 - Nov. 2025

- Researched and developed counterfactual video generation pipelines to reduce hallucinations in **Video Language Models (VLMs)**.
- Engineered a hybrid reasoning module combining VLM logic with **Video Diffusion Models** and image editing priors.
- **Outcome:** Selected as the **inaugural EMEA recipient** of the US-based Post-Internship Research Fellowship, establishing a funded research collaboration to continue this work.

### AlmageLab - University of Modena and Reggio Emilia

Modena, Italy

PHD RESEARCHER

Nov. 2023 - Present

- Developing novel architectures for **Responsible Multimodal AI**, leading to publications at top-tier venues (**CVPR Highlight**, **ECCV**).
- Presented the **VISU** dataset and fine-tuned open-source models (Llama/CLIP) to mitigate unsafe content generation and retrieval.
- Leading a joint research track with **Amazon Prime Video** on AI Safety, bridging academic novelty with industrial scale.

Experis s.r.l.

LECTURER - MASTER IN AI & ML

- Delivered advanced lectures on Deep Learning and Computer Vision for the "Smart Factory" Master's program.

Modena, Italy

Jan. 2025 - Mar. 2025

Hipert Lab - University of Modena and Reggio Emilia

RESEARCH FELLOW

- Implemented geometric deep learning methodologies for **6D Pose Estimation** in autonomous driving contexts.
- Optimized data-driven machine learning pipelines for real-time object detection.

Modena, Italy

Feb. 2022 - Aug. 2022

Hipert Lab

RESEARCH FELLOW INTERN

- Bachelor Thesis: Applied deep learning tools for labeling and 3D pose estimation of road objects.

Mantova, Italy

May 2021 - Sep. 2021

Technical Skills

Research Areas	Generative AI, Vision-Language Models (VLMs), AI Safety & Alignment, Geometric Deep Learning, NLP
Frameworks	PyTorch (Advanced), TensorFlow, JAX (Familiar), NumPy, Pandas, Scikit-learn, OpenCV, HuggingFace
Infrastructure	Slurm Workload Manager, Docker, Git, Linux/Unix Environments, AWS EC2
Languages	Python, C++, C, Java, Bash, LaTeX, SQL

Academic Service

Conference Reviewer	ARR (2026), CVPR (2025), ECCV (2024), CVPRW (2024)
Journal Reviewer	IEEE TPAMI (2025), Pattern Recognition (2024), CVIU (2024)

Research Grants & Projects

Amazon Post-Internship Research Fellowship

Seattle, USA / Modena, Italy

PRINCIPAL INVESTIGATOR (STUDENT)

2025 - 2026

- Secured a competitive **\$30,000 research grant** (\$20K gift + \$10K compute) to fund novel research on AI Safety.
- Inaugural EMEA Recipient:** Selected to establish a direct research collaboration between Amazon US Science and Unimore.

European & National Funded Projects

Europe / Italy

RESEARCH CONTRIBUTOR

2023 - Present

- ELSA** (European Lighthouse on Safe and Secure AI): Developing guarantees for robustness in deep learning (EU Horizon).
- ELIAS** (European Lighthouse of AI for Sustainability): Contributing to resource-efficient generative models (EU Horizon).
- FAIR** (Future AI Research): National PNRR project focused on foundational multimodal research.
- ELLIS Network:** Active contributor within the ELLIS (European Laboratory for Learning and Intelligent Systems) excellence network via the ELLIOT project.

Languages

Italian	Native
English	Full Professional Proficiency