

Mirko Casu

AI Researcher · Licensed Clinical Psychologist

Catania, Italy · mirko.casu@live.it · mirkocasu.github.io · [Google Scholar](https://scholar.google.com/citations?user=0000-0001-6975-2241) · [ORCID 0000-0001-6975-2241](https://orcid.org/0000-0001-6975-2241) · github.com/mirkocasu · huggingface.co/m1rkocasu · linkedin.com/in/m1rkocasu

PROFILE

AI researcher and licensed clinical psychologist, finishing a PhD in Computer Science at the University of Catania. Builds and evaluates AI systems that observe, imitate or judge human behaviour: bias in LLM judges, synthetic-media forensics (NeurIPS 2026), and wearable and conversational AI for mental health. Currently AI research consultant at Eternal Tech; visiting researcher at EPFL in 2025–26; first-author work referenced by INTERPOL and the UK Parliament.

EXPERIENCE

AI Research Consultant · Eternal Tech, Acireale, Italy Sep 2026 – Present

Company building AI infrastructure for digital identity, persistent memory and intelligent agents

- Design the evaluation methodology and criteria for a proprietary AI technology, and build the benchmarking system that applies them
- Analyse the technology's architecture and write the scientific manuscript reporting its evaluation

Visiting Doctoral Researcher · MINTS Lab, EPFL, Lausanne, Switzerland Oct 2025 – Mar 2026

Host Prof. Andrea Cavallaro

- Ran a controlled comparison of **diffusion and autoregressive LLMs as automated judges**, characterising their evaluation biases; submitted to EACL 2027
- Analysed conversational trajectories between people and AI assistants

PhD Researcher · IPLab, University of Catania, Catania, Italy Aug 2023 – Jul 2026

SAMOTHRACE project · AI for wearables and cognitive bias in forensics

- Built **ORION**, an end-to-end pipeline that detects smoking episodes from wrist motion sensors (manuscript under review)
- Introduced the **impostor bias**, a systematic doubt that real media are authentic, and validated it with 746 participants whose accuracy at spotting AI images sat at chance (47–55.5%)
- Co-built two forensic benchmarks, **WILD** (20,000 images, 20 generators) and **SynthForensics** (6,815 videos, 5 text-to-video models; NeurIPS 2026), on which state-of-the-art detectors lose 29 AUC points on average
- Co-developed an **on-device autism screening** pipeline from children's gaze, distilled into a 1.6M-parameter ONNX ensemble that runs offline in an open-source Android app
- Led scoping reviews on AI chatbots for mental health (300+ citations, cited in UK Parliament POSTnote 737), wearables for smoking detection (49 studies, *IEEE JBHI*) and AI for eating disorders (7,545 records ranked by an LLM)

SELECTED PUBLICATIONS

SynthForensics: Benchmarking and Evaluating People-Centric Synthetic Video Deepfakes

R. Leotta, S. A. Sambataro, C. V. Ragaglia, **M. Casu**, Y. Petralia, F. Guarnera, L. Guarnera, S. Battiato · In *Advances in Neural Information Processing Systems (NeurIPS), Evaluations & Datasets Track*, 2026 · **Accepted, to appear**

Protocol-Aware, On-Device Per-Image ASD Classification on Saliency4ASD via Multi-Stream Distillation

M. O. Spata, **M. Casu**, F. Rundo, S. Battiato · *IEEE Open Journal of the Computer Society*, 2026

A (Mid)Journey Through Reality: Assessing Accuracy, Impostor Bias, and Automation Bias in Human Detection of AI-Generated Images

M. Casu, L. Guarnera, I. Zangara, P. Caponnetto, S. Battiato · *Human Behavior and Emerging Technologies*, 2025
746 participants; [open data and code](#)

WILD: a new in-the-Wild Image Linkage Dataset for synthetic image attribution

P. Bongini, S. Mandelli, A. Montibeller, **M. Casu**, et al. (17 authors) · In *2025 International Joint Conference on Neural Networks (IJCNN)*, 2025

AI Chatbots for Mental Health: A Scoping Review of Effectiveness, Feasibility, and Applications

M. Casu, S. Triscari, S. Battiato, L. Guarnera, P. Caponnetto · *Applied Sciences*, 2024

Cited in UK Parliament POSTnote 737; most-cited paper (300+ citations)

GenAI mirage: The impostor bias and the deepfake detection challenge in the era of artificial illusions

M. Casu, L. Guarnera, P. Caponnetto, S. Battiato · *Forensic Science International: Digital Investigation*, 2024

Introduced the impostor bias; referenced by INTERPOL (Project SynthWave)

29 publications in total, with PDFs and code, at mirkocasu.github.io/publications

OPEN MODELS AND SOFTWARE

Apertus 1.5 8B for Apple silicon: 12 MLX quantizations (text and vision) of the Swiss open LLM, benchmarked for perplexity and psychological profile (PsychoBench); 1,900+ downloads in the last month 2026

LLaDA 2 diffusion LLMs: 4 quantizations of LLaDA 2.1-flash and 2.2-mini, in MLX for Apple silicon and NVFP4 for NVIDIA GPUs 2026

ASD Screening Android app: Open-source app, ONNX models and training code for the IEEE OJCS paper (Apache-2.0) 2026

A (Mid)journey dataset: 75 stimuli and 11,170 human judgements, with Grad-CAM maps 2025

SKILLS

Machine learning: Python, PyTorch, Hugging Face, MLX, scikit-learn, ONNX Runtime, LLM evaluation, quantization (MLX, GGUF, NVFP4), knowledge distillation, explainability (Grad-CAM)

Human research: experimental design, mixed methods, questionnaire design, inferential statistics (Jamovi), PRISMA scoping reviews, LLM-assisted screening

Clinical and devices: psychological assessment, virtual reality, wearables, mHealth, Android

Languages: Italian (native) · English (C1, IELTS Academic 7.0) · French (basic)

EDUCATION

PhD, Computer Science · University of Catania · defence pending 2023 – 2026
Thesis: *Human Behaviour as Target and Medium of Artificial Intelligence in Clinical and Forensic Contexts*

MSc, Clinical Psychology · University of Catania 2020 – 2023

BSc, Psychological Sciences · University of Sassari 2015 – 2019

Licensed Psychologist · Order of Psychologists of Sicily, Section A, reg. no. 12260 since Sep 2024

SERVICE AND TALKS

Session Chair, Technical Session 6, IEEE ICAISF 2026, Catania 2026

Track Co-Chair, DFF-2025 Deepfake Forensics Workshop, ACM Multimedia 2025 2025

Dataset Committee, AADD-2025 challenge (13 teams), ACM Multimedia 2025 2025

Lead Organizer, FHDA challenge, IEEE MetroXRINE 2025 2025

6 talks and posters, including the Global Forum on Nicotine and IJCNN 2025 (VERIMEDIA) 2024 – 2025