

Mihaela Cătălina Stoian

Department of Computer Science – University of Oxford – Oxford, UK
✉ mihaela.stoian@cs.ox.ac.uk • 🌐 mihaela-stoian.github.io

Research Interests

Neuro-Symbolic AI • Knowledge-Aided Machine Learning • Deep Generative Modelling • Large Language Models

Education

University of Oxford

DPhil in Computer Science

Oxford, UK

2021–2026

- **Thesis:** Knowledge-Refined Deep Neural Networks for Real-World Applications
- **Supervisor:** Prof. Thomas Lukasiewicz
- Developed the **first frameworks** that make deep generative models for tabular data provably compliant with background knowledge — expressed as linear constraints and disjunctions of linear inequalities capturing non-convex and even disconnected feasible spaces.
- **Areas:** deep learning, generative modelling, autonomous driving, neuro-symbolic AI.

The University of Edinburgh

MInf (Master of Informatics), First Class Honours

Edinburgh, UK

2014–2019

- **Master's thesis:** Improving extrinsic performance in zero-resource settings
- **Bachelor's thesis:** Multilingual bottleneck features for zero-resource speech processing
- **Supervisor:** Prof. Sharon Goldwater
- **Areas:** NLP, speech-to-text machine translation, speech processing.

Selected Publications

- [1] J. Ong Jun Leang, Y. Zhao, **M. C. Stoian**, W. Li, S. B. Cohen, E. Giunchiglia. Can I Have Your Order? Monte-Carlo Tree Search for Slot Filling Ordering in Diffusion Language Models. *ICML* 2026.
- [2] **M. C. Stoian**, T. Lukasiewicz, E. Giunchiglia. A Survey on Deep Learning Approaches for Tabular Data Generation: Utility, Alignment, Fidelity, Privacy, Diversity, and Beyond. *TMLR* 2026.
- [3] **M. C. Stoian**, E. Giunchiglia. Beyond the Convexity Assumption: Realistic Tabular Data Generation under Quantifier-Free Real Linear Constraints. *ICLR* 2025.
- [4] **M. C. Stoian**, S. Dyrnishi, M. Cordy, T. Lukasiewicz, E. Giunchiglia. How Realistic Is Your Synthetic Data? Constraining Deep Generative Models for Tabular Data. *ICLR* 2024.
- [5] **M. C. Stoian**. Deep Learning with Requirements in the Real World. *IJCAI* Doctoral Consortium 2024.
- [6] **M. C. Stoian**, A. Tatomir, T. Lukasiewicz, E. Giunchiglia. PiShield: A PyTorch Package for Learning with Requirements. *IJCAI* 2024.
- [7] E. Giunchiglia, **M. C. Stoian**, S. Khan, F. Cuzzolin, T. Lukasiewicz. ROAD-R: The Autonomous Driving Dataset with Logical Requirements. *Machine Learning* 112, 2023.

Work Experience

Imperial College London

Research Associate

London, UK

2025–present

- **PI:** Dr. Eleonora Giunchiglia
- **Autoformalisation:** building methods to constrain language models so they translate informal mathematics into formal statements verifiable by proof assistants such as Lean.
- Funded by **Renaissance Philanthropy**.

FiveAI

Research Intern — Computer Vision & Autonomous Vehicles

Oxford, UK

2020–2021

- **Supervisor:** Dr. Tommaso Cavallari
- **3D symmetry detection:** designed a method that estimates planar reflective symmetries on 3D inputs by slicing along the height dimension and feeding slices to a 2D convolutional recurrent regressor, avoiding costly 3D convolutions.
- **Patent:** *3D Perception*, US 20240212189 A1 (Five AI Ltd, 2024; pending).

The University of Edinburgh
Research Assistant — NLP & Machine Translation

Edinburgh, UK
2019

- **Supervisor:** Prof. Sharon Goldwater
- **Speech-to-text translation:** showed that ASR pretraining with target-language data augmentation improves translation performance, and that pretrained-ASR word error rate directly predicts translation quality.
- Published at ICASSP 2020.

ETH Zurich
Summer Research Fellow — Network Verification & Program Synthesis

Zurich, Switzerland
2018

- **Supervisors:** Prof. Martin Vechev, Asst. Prof. Dana Drachler Cohen
- **Program synthesis for P4:** built a Python parser for P4 (Programming Protocol-Independent Packet Processors) and used the Z3 solver to encode and synthesise intended packet-processing behaviour.

Software & Patents

Software: PiShield — a PyTorch framework for knowledge-aided ML [github.com/mihaela-stoian/PiShield]

Patent: *3D Perception*, US 20240212189 A1 (Five AI Ltd, 2024; pending)

Awards

IX-WAI Early Career Grants , Imperial College London	2026
Oxford PhD Runner-up Prize , G-Research	2025
EPSRC Scholarship for Doctoral Studies , University of Oxford	2021–2025
Women in Quant Finance Grant , G-Research	2025
IJCAI Grant , IJCAI-AIJ	2024
Conference Travel Grants , St Hilda's College, Oxford	2023, 2024
Best Paper Award , AI4AD Workshop @ IJCAI	2022

Selected Talks

OxBridge Women in CS Conference , University of Oxford (<i>selected for oral presentation</i>)	2026
Civil Service Leadership Group , Imperial College London (<i>invitation-only</i>) <i>One of only four projects selected; presented to senior UK civil servants, incl. the Head of the Civil Service.</i>	2025
Dagstuhl Seminar on Logic and Neural Networks (<i>invitation-only</i>)	2025
IJCAI Doctoral Consortium	2024

Technical Skills

Languages — Python (*proficient*); C++, Java, Bash/Shell (*familiar*)

Deep learning — PyTorch, PyTorch Lightning, TensorFlow; Hugging Face, verl

ML & data — scikit-learn, NumPy, SciPy, Pandas, Matplotlib, Seaborn; Jupyter

Tooling — Linux/command line, Git, Docker, Slurm, Conda, Weights & Biases, TensorBoard, $\LaTeX$

Mentoring & Teaching

Tutorial: Beyond Soft Penalties: Hard Constraints for Neural Networks with Continuous Outputs — UAI 2026 (*co-presented*)

Co-supervision: A. Godun (MSc, TU Wien); S. Lee (MSc, Imperial); L. Pejic (BSc, TU Wien)

Teaching: Class Tutor, University of Edinburgh (2017–2019) — *Discrete Mathematics; Algorithms & Data Structures; Formal & Natural Languages*

Leadership & Service

Organiser: ROAD++ Workshop & Challenge (ECCV 2024, ICCV 2023); ROAD-R Challenge (NeurIPS 2023)

Reviewing: NeurIPS, ICLR, ICML, IJCAI, NeSy; *Machine Learning* journal

Languages

English (*proficient*) • Romanian (*native*) • German (*elementary*)