



2nd International Workshop Forging Trust in Artificial Intelligence @ IJCNN 2025 5 July 2025 | Rome, Italy

Trilussa room, Pontificia Università Gregoriana

Program

10:00-10:05

Workshop opening

Katerina Tzompanaki and Corina Besliu (ForgtAI workshop chairs)

Session 1

10:05-10:50

Invited Talk: Towards Explainable-by-design Detection of Bugs in IID and Sequence Datasets

Vassilis Christophides, Professor, Ecole Nationale Supérieure de l'Electronique et de ses Applications (ENSEA) and ETIS lab, France

10:50-11:10

Regular Paper: UGCE: User-Guided Incremental Counterfactual Exploration

Christos Fragkathoulas and Evaggelia Pitoura

11:10-11:30

Regular Paper: Cascade Adversarial Attack Search

Ziwen Wang, Daoli Shen, Xiangkun Sun, Shangshang Yang, Xiaoshan Yu, and Ye Tian

11:30-11:40

Break

Session 2

11:40-12:25

Invited Talk: Trustworthiness in AI and Autonomous Systems

Anastasios Tefas, Professor, Aristotle University of Thessaloniki, Greece

12:25-12:35

Short Paper: What should GAN in AI stand for?

Dipankar Dasgupta, Arunava Roy

12:35-12:45

Short Paper: A conformal prediction framework for natural language ambiguity detection in surgical robot instructions

Ana Davila, Jacinto Colan, Yasuhisa Hasegawa

12:45-12:55

Short Paper: Trustworthy High-Resolution Pose Estimation Through Active Perception Using Vision Transformers

Emmanouil Eleftheriadis, Theodoros Manousis, Nikolaos Passalis, and Anastasios Tefas

12:55-13:00

Workshop Closing

Katerina Tzompanaki and Corina Besliu

Sponsored by:

Bevia LTD (<https://bevia.eu>) & **PANDORA European Project** (<https://pandora-heu.eu/>)