

2023 IEEE International Conference on Artificial Intelligence Testing (AITest 2023)

**Athens, Greece
17-20 July 2023**



**IEEE Catalog Number: CFP23S64-POD
ISBN: 979-8-3503-3630-6**

**Copyright © 2023 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP23S64-POD
ISBN (Print-On-Demand):	979-8-3503-3630-6
ISBN (Online):	979-8-3503-3629-0
ISSN:	2835-3552

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2023 IEEE International Conference On Artificial Intelligence Testing (AITest) **AITest 2023**

Table of Contents

Message from the IEEE CISOSE 2023 General Chairs	ix
Message from the AITest 2023 General Chairs	x
Message from the AITest PC Chairs	xi
AITest2023 Organizers	xii
AITest2023 Program Committee	xiv

Session 1

A Classification Study on Testing and Verification of AI-Based Systems	1
<i>Emanuele De Angelis (IASI-CNR, Italy), Guglielmo De Angelis (IASI-CNR, Italy), and Maurizio Proietti (IASI-CNR, Italy)</i>	
A Trustworthiness Score to Evaluate DNN Predictions	9
<i>Abanoub Ghobrial (University of Bristol, UK), Darryl Hond (RTI; Thales UK, UK), Hamid Asgari (RTI; Thales UK, UK), and Kerstin Eder (University of Bristol, UK)</i>	
DEANOMALYZER: Improving Determinism and Consistency in Anomaly Detection Implementations	17
<i>Muyeed Ahmed (New Jersey Institute of Technology, USA) and Iulian Neamtii (New Jersey Institute of Technology, USA)</i>	

Session 2

Uncertainty Quantification for Machine Learning Output Assurance Using Anomaly-Based Dataset Dissimilarity Measures	26
<i>Gabriele Incorvaia (Thales UK, Research, Technology & Innovation, UK), Darryl Hond (Thales UK, Research, Technology & Innovation, UK), and Hamid Asgari (Thales UK, Research, Technology & Innovation, UK)</i>	
Metamorphic Testing For Machine Learning: Applicability, Challenges, and Research Opportunities	34
<i>Faqeer ur Rehman (Montana State University, USA) and Madhusudan Srinivasan (University of Nebraska, USA)</i>	
Difficulty and Severity-Oriented Metrics for Test Prioritization in Deep Learning Systems	40
<i>Hamzah Al-Qadasi (Univ. Grenoble Alpes, France), Yliès Falcone (Univ. Grenoble Alpes, France), and Saddek Bensalem (Univ. Grenoble Alpes, France)</i>	

Session 3

Fault Localization for Reinforcement Learning	49
<i>Jesús Morán (University Of Oviedo, Spain), Antonia Bertolino (Consiglio Nazionale delle Ricerche, Italy), Claudio de la Riva (University Of Oviedo, Spain), and Javier Tuya (University Of Oviedo, Spain)</i>	
A Testing Framework for AI Linguistic Systems (testFAILS)	51
<i>Y. Kumar (Kean University, USA), P. Morreale (Kean University, USA), P. Sorial (Kean University, USA), J. Delgado (Kean University, USA), J. Jenny Li (Kean University, USA), and P. Martins (Kean University, USA)</i>	
AI Decision Assistant ChatBot for Software Release Planning and Optimized Resource Allocation	55
<i>Ehab Ebrahim (Ain Shams University, Egypt), Mazen Sayed (Ain Shams University, Egypt), Mahmoud Youssef (Ain Shams University, Egypt), Hisham Essam (Ain Shams University, Egypt), Salma Abd El-Fattah (Ain Shams University, Egypt), Dina Ashraf (Ain Shams University, Egypt), Omar Magdy (Siemens Digital Industries Software, Egypt), and Reem ElAdawi (Siemens Digital Industries Software, Egypt)</i>	
Visualization Tool for Extraction of Various Attributes and Corresponding Data for Dataset Quality Assessment	61
<i>Masatoshi Sekine (Hitachi, Ltd., Japan), Daisuke Shimbara (Hitachi, Ltd., Japan), Tomoyuki Myojin (Hitachi, Ltd., Japan), and Eri Imatani (Hitachi, Ltd., Japan)</i>	

Session 4

A Framework for Autonomous Vehicle Testing Using Semantic Models	66
<i>Yejun He (Nanjing University of Aeronautics and Astronautics, China), Muslim Razi (San Jose State University, USA), Jerry Gao (San Jose State University, USA), and Chuanqi Tao (Nanjing University of Aeronautics and Astronautics, China)</i>	
Evaluation of Out-of-Distribution Detection Performance on Autonomous Driving Datasets	74
<i>Jens Henriksson (Semcon AB, Sweden), Christian Berger (University of Gothenburg, Sweden), Stig Ursing (Semcon AB, Sweden), and Markus Borg (Lund University, Sweden)</i>	
Robustness Testing of an Industrial Road Object Detection System	82
<i>Anne-Laure Wozniak (Kereval; ENSTA Bretagne, France), Ngoc Q. K. Duong (Lacroix Impulse, France), Ian Benderitter (Lacroix Impulse, France), Sarah Leroy (Kereval, France), Sergio Segura (Universidad de Sevilla, Spain), and Raúl Mazo (ENSTA Bretagne, France)</i>	

Session 5

Time Series Anomaly Detection Using Convolutional Neural Networks in the Manufacturing Process of RAN	90
<i>Cristina Landin (Örebro university, Sweden), Jie Liu (Ericsson AB, Sweden; Technical University of Berlin, Germany), Katerina Katsarou (Technical university of Berlin, Germany), and Sahar Tahvili (Ericsson AB, Sweden)</i>	
TCP-Net++: Test Case Prioritization Using End-to-End Deep Neural Networks - Deployment Analysis and Enhancements	99
<i>Mohamed Abdelkarim (Siemens Digital Industries Software, Egypt) and Reem ElAdawi (Siemens Digital Industries Software, Egypt)</i>	
Leveraging Deep Learning Models for Cross-Function Null Pointer Risks Detection	107
<i>Yue Ding (Baidu), Qian Wu (Baidu), Yinzhu Li (Baidu), Dongdong Wang (Baidu), and Jiaxin Huang (Baidu)</i>	

Session 6

Using Neural Networks for Novelty-Based Test Selection to Accelerate Functional Coverage Closure	114
<i>Xuan Zheng (University of Bristol, UK), Kerstin Eder (University of Bristol, UK), and Tim Blackmore (SiFive, UK)</i>	
Unlocking the Potential of Dynamic Languages: An Exploration of Automated Unit Test Generation Techniques	122
<i>Manuel Magalhães (Universidade da Beira Interior, Portugal), Ângelo Morgado (Universidade da Beira Interior, Portugal), Henrique Jesus (Universidade da Beira Interior, Portugal), and Nuno Pombo (Universidade da Beira Interior, Portugal)</i>	
Evaluating Interdisciplinary Concept Association Methods Based on Representation Learning	N/A
<i>Zhongyi Wang (Central China Normal University, China), Guangzhao Zhang (Central China Normal University, China), Haoxuan Zhang (Central China Normal University, China), Jing Huang (Wuhan Polytechnic, China), and Haihua Chen (University of North Texas, USA)</i>	
Resource-Aware Online Permanent Fault Detection Mechanism for Streaming Convolution Engine in Edge AI Accelerators	132
<i>Weison Lin (The University of Edinburgh, UK), Xianpo Ni (The University of Edinburgh, UK), and Tughrul Arslan (The University of Edinburgh, UK)</i>	

Session 7

Enhancing Text Classification Models with Generative AI-Aided Data Augmentation	138
<i>Huanhuan Zhao (The University of Tennessee, USA), Haihua Chen (The University of North Texas, USA), and Hong-Jun Yoon (Oak Ridge National Laboratory, USA)</i>	

Measurement and Identification of Informative Reviews for Automated Summarization	146
<i>Huyen Nguyen (University of North Texas, USA), Haihua Chen (University of North Texas, USA), Roopesh Maganti (University of North Texas, USA), KSM Tozammel Hossain (University of North Texas, USA), and Junhua Ding (University of North Texas, USA)</i>	

Session 8

Debunk Online Rumors via Generative-Contrastive Learning Based on tree-Transformer	152
<i>Xi Luo (BNU-HKBU United International College; Hong Kong Baptist University, China), Yuhui Deng (BNU-HKBU United International College, China), Junjie Liu (BNU-HKBU United International College; Hong Kong Baptist University, China), Sirong Wu (BNU-HKBU United International College; Hong Kong Baptist University, China), and Gengchen Sun (The University of Hong Kong, China)</i>	
Evaluating the Impact of Incentive/Non-Incentive Reviews on Customer Decision-Making	160
<i>Kate Kargozari (University of North Texas, USA), Junhua Ding (University of North Texas, USA), and Haihua Chen (University of North Texas, USA)</i>	
Leveraging Cognitive Science for Testing Large Language Models	169
<i>Ramya Srinivasan (Fujitsu Research of America), Hiroya Inakoshi (Fujitsu Research), and Kanji Uchino (Fujitsu Research of America)</i>	
Author Index	173