AICompS 2025

The International Conference on Artificial Intelligence Computing and Systems



November 26 - 28, 2025 Fukuoka, Japan & Hybrid

Program

11:30-13:00 13:00*13:50 13:50*14:10 14:10-15:10 15:40-17:00 15:40-17:00 18:00-21:00 Nov 28 09:30-10:50 11:20-12:20 11:20-12:20 14:00-14:50	Opening Rema 15:00–16:00: Poster Stage Presentation (3 min. each) 16:00–16:30: Break 16:30–18:30: Presentation in front of the posters & Reception S1. LLM Evaluation & Reliability Not Just Accuracy: Consistency and Reliability as Core Factors in LLM Evaluation MiniSketch-VQA: A Benchmark for Evaluating LLM-as-a-Judge in Sketch-based Visual Question Answering The Dual-Edged Sword of Instruction Tuning: An Empirical Study on Precision Gains and Catastrophic Failures Evaluating Emoji Sequence Generation in Small Language Models "Hyperscale Bug Finding: "S3. Privacy in Federated / On-Device ML Mitigating Label Inference Attacks in Vertical Federated Split Learning through Training Control and Gradient Disturbance Data Reconstruction Attacks against Privacy Preserving Vertical Federated Learning ARS-FL-IDS: Accountable Anonymous Federated Learning Against	a \$4M grand prize in the e augmented large lan	S2. Prompt Safety & Grounding Vague2Detect: Handling Ambiguous Prompts in Knowledge-Based Open-World Detection PRISM: Prompt Risk Scoring via Interpretable Semantic Mapping for NSFW Defense Strong Detection, Short Text: A Four-Class Framework for Human-Al Authorship Empowering LLM-based Malware Analysis with Synthetic Code and Team Atlanta" Taesoo Kim, Professor, Georgia Tech the final round. In this talk, I will introduce the DARPA AlxCC competition in guage models (LLMs) with traditional software analysis techniques to rojects.	Kouichi Sakura (Kyushu Univ.) Chansu Han (NICT) Tao Ban (NICT) Taekyoung Kwo (SNU)
15:00-18:30 1 1 1 1 1 1 1 1 1	Opening Rema 15:00–16:00: Poster Stage Presentation (3 min. each) 16:00–16:30: Break 16:30–18:30: Presentation in front of the posters & Reception S1. LLM Evaluation & Reliability Not Just Accuracy: Consistency and Reliability as Core Factors in LLM Evaluation MiniSketch-VQA: A Benchmark for Evaluating LLM-as-a-Judge in Sketch-based Visual Question Answering The Dual-Edged Sword of Instruction Tuning: An Empirical Study on Precision Gains and Catastrophic Failures Evaluating Emoji Sequence Generation in Small Language Models "Hyperscale Bug Finding: "S3. Privacy in Federated / On-Device ML Mitigating Label Inference Attacks in Vertical Federated Split Learning through Training Control and Gradient Disturbance Data Reconstruction Attacks against Privacy Preserving Vertical Federated Learning ARS-FL-IDS: Accountable Anonymous Federated Learning Against	Jihie Kim (Dongguk Univ.) Lun Keynote I and Fixing: DAPRA 9 \$4M grand prize in the e augmented large lan e-scale open-source pr Coffee Akira Otsuka (IISEC), Tao Ban	S2. Prompt Safety & Grounding Vague2Detect: Handling Ambiguous Prompts in Knowledge-Based Open-World Detection PRISM: Prompt Risk Scoring via Interpretable Semantic Mapping for NSFW Defense Strong Detection, Short Text: A Four-Class Framework for Human-Al Authorship Empowering LLM-based Malware Analysis with Synthetic Code her. AlxCC and Team Atlanta" Taesoo Kim, Professor, Georgia Tech her final round. In this talk, I will introduce the DARPA AlxCC competition inguage models (LLMs) with traditional software analysis techniques to rojects. Break	(Kyushu Univ.) Chansu Han (NICT) Tao Ban (NICT) Taekyoung Kwo (SNU) Brent ByungHoo Kang
11:30-13:00 13:00-13:50 13:50-14:10 14:10-15:10 15:40-17:00 15:40-17:00 18:00-21:00 11:20-12:20 11:20-12:20 11:20-12:20 14:50-14:50	15:00—16:00: Poster Stage Presentation (3 min. each) 16:00—16:30: Break 16:30—18:30: Presentation in front of the posters & Reception S1. LLM Evaluation & Reliability Not Just Accuracy: Consistency and Reliability as Core Factors in LLM Evaluation MiniSketch-VQA: A Benchmark for Evaluating LLM-as-a-Judge in Sketch-based Visual Question Answering The Dual-Edged Sword of Instruction Tuning: An Empirical Study on Precision Gains and Catastrophic Failures Evaluating Emoji Sequence Generation in Small Language Models "Hyperscale Bug Finding." Team Atlanta placed 1st in the DARPA Al Cyber Challenge (AlxCC), earning and share our technical approaches that led to victory—specifically, how wautomatically discover and repair security vulnerabilities in real-world, larg S3. Privacy in Federated / On-Device ML Mitigating Label Inference Attacks in Vertical Federated Split Learning through Training Control and Gradient Disturbance Data Reconstruction Attacks against Privacy Preserving Vertical Federated Learning ARS-FL-IDS: Accountable Anonymous Federated Learning Against	Jihie Kim (Dongguk Univ.) Lun Keynote I and Fixing: DAPRA 9 \$4M grand prize in the augmented large lan e-scale open-source pr Coffee Akira Otsuka (IISEC), Tao Ban	S2. Prompt Safety & Grounding Vague2Detect: Handling Ambiguous Prompts in Knowledge-Based Open-World Detection PRISM: Prompt Risk Scoring via Interpretable Semantic Mapping for NSFW Defense Strong Detection, Short Text: A Four-Class Framework for Human-Al Authorship Empowering LLM-based Malware Analysis with Synthetic Code och ALXCC and Team Atlanta Taesoo Kim, Professor, Georgia Tech the final round. In this talk, I will introduce the DARPA AIxCC competition in guage models (LLMs) with traditional software analysis techniques to rojects. Break	(Kyushu Univ.) Chansu Han (NICT) Tao Ban (NICT) Taekyoung Kwo (SNU) Brent ByungHoo Kang
11:30-13:00 13:00*13:50 13:00*13:50 13:50*14:10 14:10-15:10 15:40-17:00 15:40-17:00 10:50-11:20 11:20-12:20 11:20-12:20 14:50-14:50	Not Just Accuracy: Consistency and Reliability as Core Factors in LLM Evaluation MiniSketch-VQA: A Benchmark for Evaluating LLM-as-a-Judge in Sketch-based Visual Question Answering The Dual-Edged Sword of Instruction Tuning: An Empirical Study on Precision Gains and Catastrophic Failures Evaluating Emoji Sequence Generation in Small Language Models "Hyperscale Bug Finding." "Hyperscale Bug Fin	Keynote I and Fixing: DAPRA s S4M grand prize in the e augmented large lan e-scale open-source pre Coffee Akira Otsuka (IISEC), Tao Ban	Vague2Detect: Handling Ambiguous Prompts in Knowledge-Based Open-World Detection PRISM: Prompt Risk Scoring via Interpretable Semantic Mapping for NSFW Defense Strong Detection, Short Text: A Four-Class Framework for Human-AI Authorship Empowering LLM-based Malware Analysis with Synthetic Code occurred to the AIXCC and Team Atlanta Taesoo Kim, Professor, Georgia Tech the final round. In this talk, I will introduce the DARPA AIXCC competition in Indianal Software analysis techniques to rojects. Break	(NICT) Taekyoung Kwo (SNU) Brent ByungHoo Kang
13:00~13:50 13:50~14:10 14:10~15:10 15:10~15:40 15:40~17:00 18:00~21:00 Nov 28 09:30~10:50 11:20~12:20 11:20~12:20 14:00~14:50 14:50~15:25	Team Atlanta placed 1st in the DARPA AI Cyber Challenge (AIxCC), earning and share our technical approaches that led to victory—specifically, how we automatically discover and repair security vulnerabilities in real-world, larg S3. Privacy in Federated / On-Device ML Mitigating Label Inference Attacks in Vertical Federated Split Learning through Training Control and Gradient Disturbance Data Reconstruction Attacks against Privacy Preserving Vertical Federated Learning ARS-FL-IDS: Accountable Anonymous Federated Learning Against	Keynote I and Fixing: DAPRA a \$4M grand prize in the e-scale open-source pr Coffee Akira Otsuka (IISEC), Tao Ban	A AIxCC and Team Atlanta" Taesoo Kim, Professor, Georgia Tech he final round. In this talk, I will introduce the DARPA AIxCC competition guage models (LLMs) with traditional software analysis techniques to rojects. Break	Kang
13:50~14:10 14:10~15:10 14:10~15:10 15:10~15:40 15:40~17:00 15:40~17:00 10:50~11:20 11:20~12:20 11:20~14:50 14:50~15:25	Team Atlanta placed 1st in the DARPA AI Cyber Challenge (AIxCC), earning and share our technical approaches that led to victory—specifically, how we automatically discover and repair security vulnerabilities in real-world, larg S3. Privacy in Federated / On-Device ML Mitigating Label Inference Attacks in Vertical Federated Split Learning through Training Control and Gradient Disturbance Data Reconstruction Attacks against Privacy Preserving Vertical Federated Learning ARS-FL-IDS: Accountable Anonymous Federated Learning Against	and Fixing: DAPRA a \$4M grand prize in the eaugmented large lane-escale open-source pr Coffee Akira Otsuka (IISEC), Tao Ban	Taesoo Kim, Professor, Georgia Tech he final round. In this talk, I will introduce the DARPA AlxCC competition guage models (LLMs) with traditional software analysis techniques to rojects. Break	Kang
14:10-15:10 15:10~15:40 15:40-17:00 18:00-21:00 18:00-21:00 10:50-11:20 11:20-12:20 14:00-14:50	Mitigating Label Inference Attacks in Vertical Federated Split Learning through Training Control and Gradient Disturbance Data Reconstruction Attacks against Privacy Preserving Vertical Federated Learning ARS-FL-IDS: Accountable Anonymous Federated Learning Against	Akira Otsuka (IISEC), Tao Ban		
15:10~15:40 15:40-17:00 18:00-21:00 18:00-21:00 10:50-11:20 11:20-12:20 11:20-14:00 14:00-14:50	Mitigating Label Inference Attacks in Vertical Federated Split Learning through Training Control and Gradient Disturbance Data Reconstruction Attacks against Privacy Preserving Vertical Federated Learning ARS-FL-IDS: Accountable Anonymous Federated Learning Against	(IISEC), Tao Ban	S4. Crypto. FHE & Secure Computation	
15:40-17:00 18:00-21:00 18:00-21:00 10:50-11:20 11:20-12:20 11:20-14:00 14:00-14:50	Malicious Behavior		An Intra-ciphertext Optimization for Efficient Multi-device Bootstrapping for HE-DNNs Efficient Evaluation of Indicator Function with Fully Homomorphic Encryption for Privacy-Preserving Embedding Efficient Batch Verifications for KZG Commitments	Yang Li (UEC)
18:00-21:00 18:00-21:00 10:50-11:20 11:20-12:20 14:00-14:50		Coffee	Break	
10:50-11:20 11:20-12:20 11:20-14:00 14:00-14:50	S5. Medical & Multimodal AI Advancing Prehypertension Screening with Explainable AI and Generative Augmentation Multimodal Pain Intensity Assessment from Physiological Signals: Window Segmentation with Cross-Attention and Temporal Modeling Multi-Resolution Speckle Priors for Scale-Aware Digital Image Correlation Enhanced Carbon Emission Prediction in IoT using Optimized Rotation-Invariant Coordinate Convolutional Neural Network for Accurate Urban Environmental Monitoring	Jonghyun Choi (SNU)	S6. Blockchain Systems & Data Infrastructure Mitigating Data Poisoning Attack in On-Device Learning Anomaly Detectors via Peer Consensus Blockchain-Based Smart Contract Revocable Bidding Scheme for European Union Emissions Trading Scheme Performance Bottleneck Analysis and Technical Debt in a Non-Standard Hyperledger Fabric CLI Gateway Architecture Blockchain-Assisted Resource Scheduling for S8-SPS in V2X Networks Service-aware Resource Management in Cloud computing for HPC workload	Akira Otsuka (IISEC)
10:50-11:20 11:20-12:20 11:20-12:20 12:20-14:00 14:00-14:50		Banquet @) Sanshiro	
11:20-12:20 12:20-14:00 14:00-14:50	S7. Systems for Efficient AI & Real-Time Calypso: A Compiler-Runtime Framework for Configurable Kernel Support in CKL-PNM Grasle: Graph-level Scheduling Language and Framework for Deep Neural Network Design and Implementation of a Timing Monitoring System for Real-Time Assurance based on AUTOSAR Timing Protection Accelerating ONNX Runtime Inference Through Tilling and IO-Binding based Model Design	Tao Ban (NICT)	S8. Applied Al for Civic & Public Safety Spatial-Temporal Graph Neural Networks for Non-Emergency Reports: A Case Study Context-Aware Safety Report Classification via Large Language Models and Dynamic Knowledge Graphs HieraText: Unsupervised Multi-Label Hierarchical Text Classification through Adaptive Clustering Web-Search-Integrated RAG for Resource-Constrained Environments: A Small-Model Approach	Sooel Son (KAIST)
12:20-14:00 14:00-14:50		Coffee	Break	
14:00-14:50 A I I I I I I I I I I I I I I I I I I	S9. LLM Security & Evaluation at Scale The Survey of Jailbreak Attacks on Large Language Models and Defenses OpenScore: An Agent-Based Framework for Automated Evaluation of Al Model Transparency BaSTion: Backdoor Style Trigger Identification Method via GANs and Style Transfer Networks	Hyoungshick Kim (Sungkyunkwan Univ.)	S10. Attacks & Systems Security H-IDE: Hardware Interleaving for Deterministic Encryption to Mitigate Ciphertext Side-Channel Attacks Model Extraction Attack Leveraging Fractal Images on Color Image Classification Tasks Applying AI Modeling Attacks as Assessment Tool for Analog PUF-Based Authentication Protocol Evaluation	Tao Ban (NICT)
## A T I I I I I I I I I I I I I I I I I I		Lun Keynote II	ch	
	Keynote II "Navigating the AI Revolution: Ensuring Security and Safety of Frontier AI" LAM Kwok Yan, Professor of Computer Science, College of Computing and Data Science, Nanyang Technological Universit, Singapore As artificial intelligence evolves from traditional machine learning to foundation models and agentic AI, society stands at a widening frontier of both opportunity and risk. This talk will examine how accelerating capabilities, emerging autonomy, and deepening societal integration have transformed AI safety and security from isolated technical issues into systemic and socio-economic priorities. It will discuss the expanding AI attack surface across data, models, and deployment pipelines, highlighting the risk of Gen AI being misused by cyber-attackers to cyber offences. This talk will also discuss defensive approaches in response to AI risks test and evaluation, red-teaming, interpretability, monitoring etc that form the backbone of trusted AI operations. Looking ahead, it will discuss risks due to the rise of agentic AI, autonomous systems capable of goal-directed behaviour and self-adaptation and the safety and security challenges this poses.			Kouichi Sakura (Kyushu Univ.
	Al being misused by cyber-attackers to cyber offences. This talk will also dis interpretability, monitoring etc that form the backbone of trusted Al opera	Coffee		
	Al being misused by cyber-attackers to cyber offences. This talk will also dis interpretability, monitoring etc that form the backbone of trusted Al opera capable of goal-directed behaviour and self-adaptation and the safety and	Daehee Jang	S12. FinTech / Crypto Analytics CollaG: Secure and Efficient Collaborative Cloud-Assisted Garbled Circuits Hybrid GCN—GRU Model for Anomaly Detection in Cryptocurrency Transactions Hall Sensor-based Ellipse Fitting for Non-Contact Ball Joint Position Estimation Design and Implementation of KI Cloud R&D Platform for HPC	Jiwon Seo (SNU)
16:35–17:15	Al being misused by cyber-attackers to cyber offences. This talk will also dis interpretability, monitoring etc that form the backbone of trusted Al opera	(Kyunghee Univ.)	Workloads	
17:15–17:45	Al being misused by cyber-attackers to cyber offences. This talk will also dis interpretability, monitoring etc that form the backbone of trusted Al opera capable of goal-directed behaviour and self-adaptation and the safety and S11. Secure Al Supply Chain & Transparency • Fine-Tuning Large Language Models for Malicious Package Detection • Automated Vulnerability Repair based on Language Agent Tree Search • MaquillAl: Generative Al for Personalized Makeup Tutorial • Variables for Free: Breaking MAYO via Valid Solutions	(Kyunghee Univ.) Best Paper Award	Workloads	Kouichi Sakura