2026 IEEE International Conference on Web Services (ICWS 2026)

Sydney, Australia
13-18 July 2026

Pages 1–777

IEEE Catalog Number: CFP26CWS-POD
ISBN: 979-8-3195-3770-6

IEEE Advancing Technology for Humanity

Copyright © 2026, IEEE

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:	CFP26CWS-POD
ISBN (Print-On-Demand):	979-8-3195-3770-6
ISBN (Online):	979-8-3195-3769-0
ISSN:	2836-3876

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

TABLE OF CONTENTS

FedTTP: Test-Time Transformer Personalisation for Federated Learning as a Service 1
George Aziz, Sajib Mistry, Aneesh Krishna

Service Integrity of LLM-Driven Edge Web Services Under Semantic Backdoor Attacks 8
Chun Yun Wang, Wei Emma Zhang, Zhigang Lu, Yan Meng, Shuchao Pang, Quan Z. Sheng

FedMoE: Efficient Federated Fine-Tuning for MoE-Based Large Language Models 20
Hanzi Mei, Jinfeng Wen, Mengwei Xu, Ao Zhou, Shangguang Wang

Higher-Order and Dual-Geometry Modeling for Blockchain Anti-Money Laundering Detection 32
Qiuyang Zhang, Mingxuan Jiang, Yuan Shui, Zhixin Li, Hongbin Zhu, Hongfeng Chai

FlexVFA: A Flexible and Verifiable Privacy-Preserving Auditing Framework for Federated Business Processes 39
Puwei Wang, Linman Li, Chenrui Li, Jinchuan Chen

TRUEdge: Faithful Edge-Aware Graph Learning for Intrusion Detection 46
Devashish Chaudhary, Sutharshan Rajasegarar, Shiva Raj Pokhrel

Agent-Based Zero-Shot Service Composition via Large Reasoning and Action Models 53
Ilche Georgievski, Marco Aiello

GLST: An Efficient and Modular Framework for Spatio-Temporal Service Flow Forecasting 60
Wen Xie, Xiaofei Xu, Jiayi Qu, Hanlin Fan, Kai Di, Shuang Yu, Mingyi Liu, Zhongjie Wang

Ranking-Aware Semantic Alignment for LLM-Based Recommendation 71
Chaopeng Dou, Yongqiang Gao, Yongmei Liu, Zijian Qiao, Jingjin Luo

Loki's Gaze: An Efficient Detection Method of LLM Backdoor with Prompt Injection Attacks 78
Hengqi Guo, Yuebing Zhou, Rongzhen Li, Qiang Duan

ChunkZ: Fast Chunked Climate Data Compression Using Spatio-Temporal Predictors 89
Shukai Chen, Chengcheng Shao, Jiaming Tan, Yazhen Wang, Xiaoyong Li, Xiang Zhu, Kaijun Ren

Auditing Black-Box LLM APIs via Token Probabilities and Data Provenance: A Case Study in Source Code Generation 96
Subhankar Maity, Faiza Belbachir, Omar Boucelma, Guilherme Medeiros Machado

AMOS: An Affinity-Aware Scheduling Method for Openharmony Services in MEC Scenario Driven by Elite-Micro-Population 103
Jie Chen, Bohai Zhao, Kai Peng, Xiaolong Xu, Zhiying Tu, Yongchao Xing, Jiaxin Liu, Dianhui Chu

From Latent Embeddings to Explainable QoS Preferences: A Generative Inference Model with Feasibility Regularization 115
Min Li, Hanchuan Xu

CompQ: QoE-Aware Compensation for Resumed Requests in Interactive LLM Inference Services 127
Yinming Tan, Jinquan Zhang, Long Chen, Kai Di, Pan Li

Fully Decentralized Mixture-Of-Experts-Based Placement of AI-Enabled Fog Applications 134
EunChan Park, In-Young Ko

A File-Centric Collaborative Migration Framework for Containerized Services in Edge Computing 146
Jiangwei Li, Zhangbing Zhou, Sami Yangui, Deng Zhao, Yixuan Chen, Zhenjie Yu

SG-LR QoS: Sparse-Gating and Layered Low-Rank Based Lightweight QoS Monitoring for Multi-Tenant Clouds 154
Pengwei Liu, Minxian Xu, Jiongjiong Gu, Chuanfei Xu, Weipeng Cao, Zhong Ming

Social Recommendation with Generative Adversarial Denoising 161
Junwei Hu, Xiantao Cai, Wenbin Hu

ACCORD: Accountable Credential Disclosure for Self-Sovereign Identity 173
Evan Krul, Hye-young Paik, Sushmita Ruj, Salil S. Kanhere

Lean Service Catalogs for Agentic Systems: A Study of Schema Compression and Caching Economics 177
Anthony Shaw, Rong N. Chang, Amin Beheshti

What Is a Data Domain? a Formal Metamodel and Canvas for Decentralized Data Architectures 184
Marco Tonnarelli, Filippo Scaramuzza, Laura Schuiki, Damian Andrew Tamburri, Indika Kumara Stefan Driessen

ALPINE: Adaptive Privacy-Budget Control for Structured Agent-to-Agent Communication in Distributed Multi-Agent Service Systems 196
Guanjie Cheng, Siyang Liu, Yishan Chen, Mengying Zhu, Hailiang Zhao, Zhen Qin, Yangyang Wu Xinkui Zhao

Fides: Federated Inter-Organizational System Dynamics for Explorative Process Simulation 208
Silas Workman, Joyce El Haddad, Daniela Grigori

Test-Time Adaptive Composition for Machine Learning as a Service (Mlaas) in IoT Environments 215
Deepak Kanneganti, Sajib Mistry, Sheik Mohammad Mostakim Fattah, Aneesh Krishna

Large or Small? Cloud or Edge? Investigating the Use of Language Models for Smart Drone Services 222
Peilin He, Xiao Liu, Jia Xu, Jonathan Kua, Disen Jia

AnchorMark: Preventing Watermark Collapse in LLMS via Anchor-Based Resynchronization 229
Qirui Sa, Yifei Zhang, Weijing You, Cunqing Ma, Yiwei Shan, Yongjuan Sun, Qi Gu, Wenmao Liu

HPRCC: Popularity-Aware and Physics-Inspired Content Caching in MEC by Using Adversarial Semi-Bandit Learning 240
Jinpeng Li, Xiaoning Sun, Houli Xie, Yunni Xia, Peng Chen, Ziyang Zeng, Lijun Liu

Hypergraph-Guided Distribution Alignment for Bundle Recommendation 251
Wenyao Qiao, Sibo Wen, Nan Wang, Yingli Zhong, Lingli Li

Uncertainty-Aware Zygote Container Planning for Burst-Resilient Function Startup in Serverless Edge-Cloud Computing 262
Tianyu Fang, Jiangwei Li, Xiaocui Li, Zhangbing Zhou

Temporal Edge-Enhanced Bipartite Graph Neural Networks for Secondary Battery Swapping Service 271
Chunqi Wu, Zhao Li, Jinsheng Ding, Pan Tang, Donghui Ding, Jiaming Huang, Chonghui Zheng Yuhua Zheng

TGHE: Template-based Graph Homomorphic Encryption for Privacy-Preserving GNN Inference in Edge-Cloud Systems 278
Ngoc Bao Anh Le, Thai T. Vu, John Le, Heath Cooper, Jun Shen

DCQ: Dimension Selection and Compensation in Post-Training Quantization for Edge Vision Services 285
Junna Zhang, Luxi Wang, Chuntao Ding, Zigui Jiang, Guoqing Zheng

ATMOcloud: Automated Threat Modeling for Cloud Architectures Through Cyber Threat Intelligence Enrichment and Structured Validation 292
Noam Tarshish, Avishag Shapira, Ido Hersko, Yuval Elovici, Asaf Shabtai

FiLA-DP: Fisher-Guided Layer Sharing for Differential Privacy in Federated Learning 299
Qiufen Ni, Shangyuan Xie, Xiao Dong, Jianxiong Guo, Tian Wang

MGLA: Multimodal Graph Learning Based Algorithm for Candidate Microservice Identification 310
Jianlin Zhang, Cong Liu, Yuliang Shi, Shijun Liu, Dong Zhang, Kaiyuan Qi, Zhiyong Chen

iRUC++: Traffic-Aware Adaptive Unified Computation for Communication-Dense Microservices 319
Xinyi Liu, Puwei Wang, Keman Huang, Xiaoyong Du

Collaborative Assurance of on-Device Model Inference Services in AIoT Environment 330
Ke Meng, Xiaolong Xu, Haolong Xiang, Peng Chen, Guangming Cui, Xuyun Zhang, Lianyong Qi, Lina Yao Wanchun Dou

CogGuard: Cognitive and Operational Profiling for Proactive Warning in Edge Intelligent Services 341
Zhi Yao, Weihao Chen, Zhiqing Tang, Hanshuai Cui, Qianli Ma, Weijia Jia, Wei Zhao

ST-RCA: A Task-Oriented Causal Framework for Robust Spatio-Temporal Root Cause Analysis in Complex IoT Systems 352
Borui Li, Lun Tang, Qianbin Chen

LLMFlowGO: An LLM-Driven System for Workflow Generation and Optimization 358
Jia Xu, Weibin Zhang, Xuejun Li, Xiao Liu, Peilin He, Jinglong Fang, Yun Yang

Optimizing FaaS Platforms for Orchestrating MCP-Enabled Agentic Workflows 369
Varad Kulkarni, Vaibhav Jha, Nikhil Reddy, Haseeb Abdulla Kollorath, Anand Eswaran Praveen Jayachandran, Yogesh Simmhan

Intent-Driven and Constraint-Verified Service Composition via Evolving Verified Experience 381
Bowen Ma, Junxian Liu, Bangbang Ren, Lailong Luo, Deke Guo, Honghui Chen

SABALANCE: Semantic-Aware Load Balancing for Stateful NFV Service Systems 391
Yuan Liu, Wendi Feng, Ao Jin

TriSA: Tri-Type Hypergraph Semantic Alignment for Metaverse Service Paradigm and Recommendation 402
Ruipeng Han, Dunlei Rong, Yeqi Zhu, Zihang Su, Xiao Wang, Hanchuan Xu

Dual-Diffusion Service Recommendation with External Knowledge Graphs 413
Junxiang Zhang, Yanting Li, Lianqi Huang, Shuo Yu, Xinshu Li, Ziqi Xu, Feng Xia

Integrating Structural and Temporal Information of Trace Cases: a Multi-Perspective Framework for Process Representation Learning 424
Dongjin Yu, Yuying Cheng, Yayu Hou, Zhongyang Li, Ke Ni, Lei Tu

CognitiveContinuum: Intent-Aware Orchestration Using LLM and Fuzzy Logic 434
Negin Akbari, Adel N. Toosi, John Grundy, Muhammad Aamir Cheema

Ada-Jelly: A Lightweight and Adaptive Subflow Sketch for High-Speed Network Monitoring 441
Zhengli Lin, Xuanwei Si, Dong Zhang, Hongyan Liu, Bohuai Xiao, Ming Li, Jianshan Zhang, Chunming Wu Xiang Chen

K3AS: QoS-Aware Kinematic 3D Asset Service with Cache-Adaptive Progressive Asset Evolution 452
Haoqi Ruan, Ruihan Hu, Zhaodong Zhang, Wenqi Wang, Zilong Hu

Tiga: Adaptive Two-Level Parallel Transaction Execution for Improving Blockchain Qos 463
Shengjie Guan, Zhen Xiao

An LLM-Enhanced Multimodal Graph Network for Causal-Aware Traffic Forecasting 475
Jiale Wu, Haolong Xiang, Xiaolong Xu, Lianyong Qi, Xuyun Zhang, Fanzhen Liu, Rong Jiang, Wei Fan

TISKE: Two-Stage Incremental and Scalable Knowledge Editing for LLM-Based Intelligent Services 486
Qiang Liu, Buqing Cao, Wenyu Zhao, Sheng Lin, Hangyu Liu, Jinjun Chen

An Ontology-Driven Service-Oriented System for ESG Metric Computation and Reporting 495
Mingqin Yu, Fethi Rabhi, Basem Suleiman, Nuodi Liu, Yuekang Li

Equivariant Positional Encoding for Graph Neural Network-Based Collaborative Filtering 502
Jishen Sun, Xin Sun, Deling Huang, Samaneh Madanian, Guiling Wang, Zhizhong Liu, Muhammad Naeem Jian Yu

Dynamic Function Deployment and Task Offloading in Serverless Multi-Edge Systems 510
Jinquan Zhang, Xiaoping Li, Zhicheng Ding, Yibo Li, Guoqiang Bao, Long Chen

TopoEvo: A Topology-Aware Self-Evolving Multi-Agent Framework for Root Cause Analysis in Microservices 521
Junle Wang, Ruibo Chen, Yuanshuai Wang, Xingchuang Liao, Zhimin Fan, Xiangdong Wu, Wenjun Wu Xuehong Liu, Jinkun Zhao

STAR: A Stage-Attributed Triage and Repair Framework for RCA Agents in Microservices 532
Junle Wang, Ruibo Chen, Yuanshuai Wang, Xingchuang Liao, Zhimin Fan, Xiangdong Wu, Wenjun Wu Jinkun Zhao, Xuehong Liu

Transferable Unsupervised Time Series Anomaly Detection with Mixture-of-Experts for Web Systems 543
Weiwei Huang, Gou Tan, Pengfei Chen

Flow Matching with Collaborative Contexts and Variational Interests for POI Recommendation 554
Yanwei Xu, Jinbiao Yao, Wenpeng Jia, Weiming Liu, Xuyun Zhang, Bin Cui

Towards Accurate Yet Efficient Federated Learning as a Service Via Personalized Lora Rank Allocation 565
Shuaijun Chen, Omid Tavallaie, Niousha Nazemi, Xin Chen, Albert Y. Zomaya

Mix and Match: Flexible De-Mixing of Bitcoin Mixers 579
Donghee Kim, Juyeon Yi, Junbeom Hur

YOLO-ADM: A Lightweight Infrared Small Target Detection Method Based on UAV Remote Sensing Images 586
Dianrui Wang, Xiangyu Bai, Haoran Cheng

Magnitude-Calibrated Contrastive Alignment for Multimodal Recommendation 593
Yacong Yang, Hang Lv, Guofang Ma, Yanchao Tan

Dual-Purpose Editable Prompts: Correcting Rating-Review Inconsistency for LLM-Based Explainable Recommendation 600
Mengying Zhu, Feiyue Chen, Mengyuan Yang, Guanjun Xu, Yangyang Wu, Guofang Ma, Guanjie Cheng Xiaolin Zheng

RISE: Relay Inference and Online Scheduling for Efficient Edge-Device Collaborative Diffusion Model Services 607
Zilan Huang, Zhiqing Tang, Hanshuai Cui, Tian Wang, Yuan Wu, Weijia Jia, Wei Zhao

Promise: Predictive and Cost-Optimal Migrations for Service-Continuous Edge Intelligence in Multi-UAV Networks 618
Jian Zhang, Ming Tao, Renping Xie, Yin Zhang

Resource-Efficient Online Reliability Anomaly Detection for Edge Services via Deep Reinforcement Learning 627
Hang Fu, Puwei Wang, Xinyi Liu, Xiaoyong Du

V-Ideal: Generative Prompting for Fault-Tolerant Multimodal Web Services Under Microservice Failures 638
Bingyan Fan, Winston K.G. Seah, Gang Xu, Xiaodong Xu, Feng Zhang

Metra: A Fault Localization Technique for Microservice Systems Via the Fusion of Metrics and Traces 648
Chang-Ai Sun, Yubo Kou, Tao Zeng, Jun Han

Swarm Network-as-a-Service (SNaaS) 660
Balsam Alkouz, Osama Amin, Basem Shihada

StockRL: Execution-Grounded Reinforcement Learning for AI-Driven Portfolio Decision Support with Dual-Level Reward Models 671
Boyang Yin, Ying Liu, Bin Zhang

Shield: Semantic Hierarchy-Enhanced Hypergraph Learning for Diagnosis Prediction 682
Xusheng Yu, Hang Lv, Guofang Ma, Yanchao Tan, Xing Chen, Carl Yang

LANCE: Linguistic Agent Network for Cooperative Ensemble of Heterogeneous Physical AI Services 689
KyeongDeok Baek, In-Young Ko

MIST: Microservices Identification from Sequential Traces 700
Thakshila Imiya Mohottige, Artem Polyvyanyy, Colin Fidge, Rajkumar Buyya, Alistair Barros

A Retrieval-Augmented Generation-Driven Auto-Scaling Method for Microservices in Cloud-Native Environments 711
Xu Jiang, Zhicheng Cai, Lei Xu, Rajkumar Buyya

SkaaS: A Skill-as-a-Service Architecture for Stateful Service Reasoning in Manual Procedures via FSM-Constrained Multimodal LLMs 722
Xinhui Li, Zhongguo Yang, Di Li, Yonglu Jiang, Juan Li

Ensuring Efficiency and Timeliness in Real-Time Rag Services: an AoI-Aware Approach 733
Jiebei Shi, Xiangru Liang, Yeping Zhu, Jianbing Zhang, Jiwei Huang

DCRL: Dynamic Causal Representation Learning for Root Cause Localization in Microservice Systems 743
Cong Liu, Jianlin Zhang, Yuliang Shi, Shijun Liu, Dong Zhang, Kaiyuan Qi, Zhiyong Chen

FMD-SQL: NL2SQL-as-a-Service for Financial Data Access 753
Xiaozheng Du, Ruijun Deng, Cheng Wang, Shijing Hu, Feng Zhou, Qiang Duan

Data Contracts: Structuring Promises and Expectations in Data Exchange 764
Laura Schuiki, Arif Wider, Simon Harrer

TrustEOV: Robust Endorsement Against Contract Manipulation in Execute-Order-Validate Blockchains 767
Zihao Guan, Bin Zhou, Jianfeng Zhang, Xiang Fu, Feng Jiang, Yan Ding

MMC-Rec: A Memory-Based Multi-Agent Coordination Framework for Recommendation 778
Xuehan Mao, Guodong Fan, Jing Li, Ming Zhu, Chun Wang

MHLAR: A Multi-View Fusion Framework Integrating Hypergraph Structures and LLM-Enhanced Semantics for Mobile App Recommendation 789
Fenfang Xie, Feng Luo, Mingdong Tang, Caijie Lin

ESM-RB: Energy-Aware Perceptual Bitrate Adaptation for Mobile Web Streaming Services 800
Wenlong Ke, Ying Wang, Shigen Shen, Yong Wang, Zhe Chen, Yangfan Liang, Junqi Chen, Yong Zhang

Toward Effective Resource Prediction for Microservices via Heterogeneity Exploration Fully Investigating Services, Containers, and Nodes 810
Yueshen Xu, Haibin Han, Shaoling Liang, Junxiao Han, Chen Zhi, Zhongxiao Wang, Huihui Tang Jiahao Wang

Self-Evolving Neural Systems for Heterogeneous Mobile Devices 821
Zhao Li, Peiqi Zhang, Enze Yu, Haipeng Dai, Qiang He, Shanyun Liu

ConCise: Training-Free Conclusion-Chain State Compression for Cost-Efficient Multi-Step RAG Services 828
Kuan Yan, Zhiqing Tang, Tian Wang, Weijia Jia

Toward Generative AI as a Service for Web Design: Efficient Multi-Style Layout Generation with Meta Networks 838
Qiwen Xiao, Lu Yu, Kai Wang, Zheng Zhang, Yuejin Li, Shijie Zhang, Xitong Li

Structgen: A Multi-Agent Generative Service for Automated Web UI Prototyping 850
Jing Luo, Shuai Zhao, Yihang Su, Jinjian Yang, Hongbo Tang, Chunguang Zhao, Yang Li

Abyss: A Service-Oriented Architecture for Web-Like Social VR 857
Minwoo Song, In-Young Ko

Dual Model Learning for Efficient and Effective Visual Services 864
Mehdi Acheli, Walid Gaaloul

FlowRAG: Control-Flow-Guided Retrieval-Augmented Generation for Service Operating Procedures 875
Congcong Tian, Zhongjie Wang, Mingyi Liu, Zhihui Xu, Xinyang Chen, Xingya Zhang, Haitao Zhao Shuang Yu

Knowledge-Grounded LLMs for Automated Business Capability Modelling from Enterprise Artefacts 886
Iromie Samarasekara, Boualem Benatallah, Madhushi Bandara, Fethi Rabhi

Neighborhood-Centric Self-Supervised Graph Models for Imputation as a Service 897
Erfan Moshiri, Jia Wu, M A Hakim Newton, Min Li, Pooyan Asgari, Amin Beheshti

PRISM: A Calibrated Multi-Signal Trust Framework for KG-Based Selective Prediction in API Service Ecosystems 907
Parimehr Morassa Far, Amin Beheshti, Kamyar Taeb, Alireza Shammasi, Ali Shahsavandi M A Hakim Newton

SCALER: LLM-Based Cross-Modal Alignment for Microservice Root Cause Analysis 918
Shu Yang, Geyuan Wang, Zhongxing Ming, Dongbiao He, Laizhong Cui, Xian Yu, Xu Zhou

TORCL: Lightweight Root Cause Localization with Temporal-Order-Aware Ranking in Microservices 929
Ruyue Xin, Peng Chen, Yunni Xia

A Two-Stage Multi-Objective Evolutionary Approach for Joint Optimization of Microservice Replication and Allocation in Energy-Aware Clouds 937
Boxiong Tan, Hui Ma, Yi Mei, Mengjie Zhang, Sven Hartmann

JPSTL: Joint Probabilistic Signal Temporal Logic for Predictive Composition Monitoring 949
Yixuan Chen, Zhangbing Zhou, Deng Zhao, Jiangwei Li

HPGPIN: A Hypergeometric Prior-Guided Physics-Informed Neural Framework for Satellite Task Scheduling 957
Wenhan Hou, Yingjie Wang, Zhaoxiang Li, Haojun Teng

PACT-IP: A Policy-Governed and Auditable Continuous IP Triage Service for SOCs 966
Di Shao, Aiting Yao, Chengzu Dong, Huan Zhang, Huchen Zhou, Zhaoquan Gu

Feedback Control-Based Self-Adaptive Cloud Resource Provisioning for Multi-Tier Web Systems 973
Yamin Lei, Jinquan Zhang, Long Chen, Zhicheng Cai, Rajkumar Buyya

OpenBinding: An Extensible, Plug-and-Play API Gateway for Cross-Engine QACO Solving 981
Francisco Javier Cavero, José Antonio Parejo, Antonio Ruiz-Cortés

QoS-Aware Semantic Routing for OGC API-EDR and Tiles Over Sharded Zarr Metocean Data Cubes 988
Jiaming Tan, Chengcheng Shao, Xiaoyong Li, Kaijun Ren, Shukai Chen, Yazhen Wang, Hui Zhang

QoS-Aware Brokering for LLM-as-a-Service: SLA-Constrained Routing with Drift Detection 999
Okba Tibermacine, Chouki Tibermacine

QoS-Aware Orchestration for on-Orbit Training Services Across LEO Satellite Systems 1008
Nan Yang, Bahman Javadi, Rodrigo N. Calheiros, David Boland, Muhammad Umair, Philip Leong

HeatMoE: A Heat-Aware Hybrid Architecture for Serverless Mixture-of-Experts Training 1019
Lingyu Song, Huiying Jin, Yunfei Zhang, Hai Dong, Pengcheng Zhang

A Novel Method to Anomaly-Adaptive Task Offloading in Vehicular Edge Computing 1030
Yujia Song, Houli Xie, Yunni Xia, Xingli Zhong, Na Zhao

FedAlloy: Dual-Channel Federated Learning Across Quantized and Heterogeneous Language Models 1041
Jianfeng Ling, Feilong Feng, Jinquan Zhang, Long Chen, Jihong Wan

TFLLM4Load: Time-Frequency Enhanced LLM-Enabled Framework for Multi-Scale Workload Prediction with External Forecasting Priors 1052
Boming Zhu, Xiaojian Gao, Chang Liu, Xiupu Lang, Haopeng Chen

Decoupled Smart Contract Audits: Lightweight LLM Framework via Distillation and Aggregation 1059
Bagus Rakadyanto Oktavianto Putra, Muhamad Risqi Utama Saputra, Guntur Dharma Putra

Early Find Early Stop: Enhancing Multi-Agent System Efficiency Via Hierarchical Real-Time Validation 1066
Zhuangyu Ma, Yanjun Shu, Longyao Zhang, Dawei Chen, Feng Xia, WeiEmma Zhang, Decheng Zuo Dongxin Wen

REKV: Cross-Precision Model Switching with Persistent KV Cache Reuse for LLM Serving 1073
Jiali Wu, Minxian Xu, Yang Wang, Kejiang Ye, Hao Dai, Chengzhong Xu

MARS: Dynamic Micro-Batching and Heterogeneous-Aware Scaling for MLLM Services in Cloud-Edge Collaboration 1080
Shikun He, Binbin Feng, Zhijun Ding

F²-AFL: A Fairness- and Freshness-Aware Approach of Asynchronous Federated Learning for IIOT 1091
Jie Chen, Jiebei Shi, Ruiqi Zhang, Yeping Zhu, Jianbing Zhang, Jiwei Huang

Agentic-AI Detection of Trust Data Tampering in Crowdsourced IoT Services 1102
Thilina Lokuruge, Athman Bouguettaya

DP-FedOR: Improving Differentially Private Federated Learning via Adaptive Origin Tracking and Residual Privatization 1113
Meiguang Zheng, Zijie Hu, Hao Zheng, Liu Yang, Xi Li, Jinlin Gui, Wenrui Li

Critical-Path-Aware Dependent Task Offloading and Scheduling for Multi-Satellite Cooperative Networks 1123
Chunhong Liu, Xiaoyu Cui, Jialei Liu, Tongguang Zhang, Bo Cheng, Peiyan Yuan

QueueBreak: A Trace-to-Diagnosis Pipeline for Tail Latency in Agentic LLM Services 1130
Xingrao Ma, Zongxi Li, Chengzu Dong, Di Shao, Xiao Liu

Minimizing Execution Delays in Configuration Management Systems via TFSM Synchronizing and Homing Sequences 1138
Evgenii Vinarskii, Natalia Kushik, Djamal Zeghlache

Detecting Vulnerability Patch Commits Via Version Filtering and Large Language Models 1145
Haoran Xu, Chen Zhi, Junxiao Han, Zhen Qin, Xinkui Zhao, Jianwei Yin, Shuiguang Deng

Multi-Source Drift-Aware Online Simulation of Evolving Business Processes 1157
Chao Zheng, Jian Wang, Bing Li

CONCORD: Asynchronous Sparse Aggregation for Device-Cloud Rag Under Document Isolation 1168
Xuedong Hu, Zhiqing Tang, Zhi Yao, Tian Wang, Weijia Jia

RE-ACTOR: Reinforced Attention Encoder for Service Conflict Resolution in IoT Systems 1178
Christson Awanyo, Nawal Guermouche

A Risk-Aware UAV-Edge Service Framework for Wildfire Monitoring and Emergency Response 1189
Yulun Huang, Zhiyu Wang, Rajkumar Buyya

QoS-Aware Soft Failure Detection for Drone Swarms via Adaptive Knowledge-Data Fusion 1200
Yi Li, Zhangbing Zhou, Yangui Sami, Alireza Furutanpey, Schahram Dustdar

Towards a Unified View of Configuration Problems in the Computing Continuum 1208
Alejandro García-Fernández, Boris Sedlak, José Antonio Parejo, Pantelis Frangoudis, Antonio Ruiz-Cortés Schahram Dustdar

Dual-Queue Lyapunov Optimization for Skill-Aware Task Assignment in Crowdsourcing 1211
Andrew Putra Kusuma, Han Yu

A Reuse-Aware Encoding-Mechanism Framework for Concurrent MEC Task Offloading 1214
Hualong Wu, Yucheng Liu, Weizhe Zhang, Desheng Wang, Hainan Zhang, Wei Gao

Towards Prosocial Equilibrium: Conformity-Driven Group Cooperation Emergence 1217
Yuheng Liu, Xiaodong Fu, Li Liu, Jie Li, Wei Peng, Jiaman Ding

Veriflood Attack: Exploiting Mandatory CAPTCHA Verification Into a Flooding Primitive 1220
Hyejin Lee, Woonghee Lee, Junbeom Hur

LLM-Driven Evolutionary Search for IoT Task Scheduling in Edge Computing 1223
Xiulei Liu, Zhenquan Liu, Yongpeng Xu

GPL: A Graph-Based Learning Approach for Mobility-Aware Multi-SFC Placement in Mobile Edge Computing 1226
Yi Yue, Ying Liu, Zhengxin Han, Shuai Zhang, Chang Cao, Xiongyan Tang

Revenue-Driven Content Deployment in Edge Communities 1229
Zhen Zhang, Hongyue Wu, Yunfei Luo, Shizhan Chen, Zhiyong Feng

Edge-Assisted Collaborative Perception Service with Correction-Detection-Prediction Architecture 1232
Yiru Lin, Ningzhe Liu, Zhijun Ding

An Adaptive Federated Service Evolution Framework for Industrial Operations Analytics Under Dynamic Participation 1235
Daobin Luo, Yuru Liu, Weishan Zhang, Yuange Liu, Qiao Qiao, Lingzhao Meng, Lianyong Qi

Generative Similar Weather Retrieval with Physics-Aligned Hierarchical Identifiers 1238
Yazhen Wang, Xiaoli Ren, Xiaoyong Li, Haibo Tang, Shukai Chen, Jiaming Tan, Kaijun Ren

Robust Privacy Unlearning for LLM Services via Adversarially Augmented Preference Optimization 1241
Shuhan Yang, Yao Guan, Qiang Duan, Chenchi Luo, Haijun Yang

EVA: An Event-Conditioned Framework for Privacy-Preserving Face Authentication Services 1244
Yuqin Zhang, Xiao Liu, Hui Zhang, Xuejun Li, Jia Xu, Frank Jiang

Royalty Distribution Models for Referable NFT-Based Services 1247
Ruiqiang Li, John Le, Brian Yecies, Qin Wang, Shiping Chen

Mitigating Semantic Confusion in LLM-Integrated Web Services via Dynamic Dual-Stream Intent Matching 1250
Zhenghao Zhao, Qingwen Zeng, Yitian Yang, Huaming Chen

Adaptive Partition and Resource Allocation for Collaborative Inference Services in UAV Swarm 1253
Ying Dong, Jiebei Shi, Songwei Li, Yuzhou Gao, Yeping Zhu, Jianbing Zhang, Jiwei Huang

Toward Value-Aware Metadata for Extreme-Event Retrieval in Cloud-Native Climate Data Services 1256
Jiaming Tan, Chengcheng Shao, Xiaoyong Li, Xiang Zhu, Kaijun Ren, Shukai Chen, Yazhen Wang

HiCoSched: Production-Trace-Driven Backtesting for LLM Online-Offline Co-Serving 1259
Chang Xu, Zhen Chen, ChuXuan Zeng, Zhiyong Zhuo, Chun Fan, Tingting Long

Towards Efficient and Reliable On-Device Multi-Agent Systems: Challenges, Technologies and Explorations 1262
Yanjun Shu, Weiye Yan, Zhuangyu Ma, Jian Yang, Yanming Jiang, Min Li, Tianyu Luo, Decheng Zuo

Towards an Agent Operating System - Lessons from Classical and Cloud OS 1274
Gosia Steinder, Hubertus Franke

Deontic Policies for Runtime Governance of Agentic AI Systems 1286
Anupam Joshi, Tim Finin, Karuna Joshi, Lalana Kagal

Optimizing Agentic AI Applications with Reinforcement Learning 1296
Federica Filippini, Hamta Sedghani, Giuseppe Vizzari, Michele Ciavotta, Danilo Ardagna

ACMA: An Efficient API Composition Multi-Agent Framework Based on Multi-Level Experience Reuse and Dynamic Strategy Selection 1301
Zehui Liu, Ying Wang, Wolfgang Mayer, Quan Z. Sheng, Pengfei Wu, Zaiwen Feng

Agentic Metaverse Services: A New As-a-Service Paradigm 1311
Xiaofei Xu, Quan Z. Sheng, Zhongjie Wang, Boualem Benatallah, Xiao Wang, Ruipeng Han

Agentic Service-Oriented Computing: A Manifesto for the Next Frontier of Service-Oriented Computing 1322
Amin Beheshti, Rong N. Chang, Boualem Benatallah, Fabio Casati, Schahram Dustdar, Geoffrey Fox Quan Z. Sheng, Yan Wang, Jian Yang, Albert Zomaya

Human Society-Inspired Approaches to Agentic AI Security: The 4C Framework 1335
Alsharif Abuadbba, Nazatul Sultan, Surya Nepal, Sanjay Jha

DualTeach: Disagreement-Aware Dual-Teacher Distillation for Reliable Agentic Services 1346
Kamyar Taeb, Wei Emma Zhang, Parimehr Morassa Far, Alireza Shammasi, Ali Shahsavandi MohammadHossein Ahmadi, Anthony Shaw, Amin Beheshti

Beyond Workflow Automation: Toward Intent-Driven Service Composition Across Heterogeneous Domains 1355
Hayet Brabra, Walid Gaaloul, Djamal Zeghlache

Intent-Driven Data Product Discovery Framework 1366
Matteo Falconi, Pierluigi Plebani

Modeling Contract-Aware Data Products: a Design Science Research at a Semiconductors Firm 1375
Jeroen Wasser, Marco Tonnarelli, Indika Kumara, Geert Monsieur, Damian Andrew Tamburri Willem-Jan van den Heuvel

LLM-Based Agent Simulation Data Product for Counterfactual Evaluation of Hotel Dynamic Pricing 1386
Mohammadali Majidi, Jian Yang, Amin Beheshti

Query-Adaptive Graph Compression for LLM-Based Service Reasoning and Composition 1397
Melika Zare, Amin Beheshti, Min Li, Pooyan Asgari

A Policy Relationship Model for Consistent Data Product Governance 1408
Valeria Maria Fortina, Pierluigi Plebani

Enabling Reuse for Data-Sharing Pipelines in Federated Environments 1419
Sepideh Masoudi, Maria C. Borges, Eduardo Brito, Sebastian Werner, Stefan Tai

LQSOA: Lyapunov-Based Queue-Stable Offloading Algorithm for Multi-UAVs in Mobile Edge Computing 1429
Yiyang Niu, Guangrui Liu, Xia Zhu

Adaptive Scheduling with Dynamic Heterogeneous Graphs for Intelligent Cloud-Edge Computing Services 1439
Jianhai Chen, Jia Wang

AI Empowering Highly Reliable Serverless Edge: A Systematic Review 1449
Yanjun Shu, Zhuangyu Ma, Tianyu Luo, Yanming Jiang, Min Li, Wei Emma Zhang

Efficient Partition-Based Subgraph Matching for Web Service 1463
Guohao Sun, Hunan Huang, Ge Zhang, Xiu Susie Fang, Jinhua Lu

QoS-Aware Prefix Orchestration for Geo-Distributed LLM-as-a-Service 1473
Zhangyang Ge, Long Chen

FTAKG: An Adaptive Fine-Tuning Multi-Agent System for Domain-Specific LLM Based on Knowledge Graph 1483
Longyao Zhang, Yanjun Shu, Hanzhi Xu, Hua Nie, Haibin Xu, Decheng Zuo, Yi Feng, Dongxin Wen

ECGNET: A Balance-Aware Multivariate Workload Predictor for Microservice Applications 1494
Yamin Lei, Tianxing Wu, Zhicheng Cai, Shuang Wang

EcoFlow: Cost-Aware LLM Serving on Heterogeneous GPUs via Min-Cost Max-Flow 1502
Fengkai Jiang, Fulu Kou, Jinquan Zhang, Rui Yang, Shujie Pang

Scaling Concurrent LLM Services via Self-Adaptive Load Balancing in Edge Environments 1513
Teng Wang, Xiang He, Shuang Yu, Zhongjie Wang

An Agent-as-a-Service Framework: A Case Study on Hydrogen Evolution 1523
Trieu Vuong Doan, Jincheol Kim, Young Choon Lee, Zhongli Dong

A Requirement-Driven Approach for Multi-Agent Service Pattern Recommendation 1529
Meng Xi, Zihao Zhang, Yechen Jin, Guanjie Cheng, Hao Guan, Ying Li, Xiaohua Pan, Jianwei Yin

Reliable Orchestration of Specialized Multi-Agents for Cloud Incident Diagnosis 1538
Haiyu Huang, Jiewei Lyu, Shuting Lai, Zhihan Jiang, Pengfei Chen, Michael R. Lyu