17th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2025)

Volume 1: KDIR

Marbella, Spain
22-24 October 2025

ISBN: 979-8-3313-3082-8

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2025) by SCITEPRESS – Science and Technology Publications, Lda.
All rights reserved.

Printed with permission by Curran Associates, Inc. (2026)

For permission requests, please contact SCITEPRESS – Science and Technology Publications, Lda.
at the address below.

SCITEPRESS – Science and Technology Publications, Lda.
Avenida de S. Francisco Xavier, Lote 7 Cv. C,
2900-616 Setúbal, Portugal

Phone: +351 265 520 185
Fax: +351 265520 186

info@scitepress.org

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
Email: curran@proceedings.com
Web: www.proceedings.com

CONTENTS

INVITED SPEAKERS

KEYNOTE SPEAKERS

Location Privacy in Crowdsourcing 5
Richard Chbeir

An Overview of AI & Law Models of Case-Based Reasoning, with Applications in Decision Analysis and Explainable AI 7
Henry Prakken

Data Quality and Ethics in Cultural Heritage Knowledge Graphs 9
Valentina Presutti

PAPERS

FULL PAPERS

Generating Aerial Flood Prediction Imagery with Generative Adversarial Networks 15
Natasha Randall, Gernot Heisenberg and Juan Luis Ramirez Duval

Sheet Metal Forming Springback Prediction Using Image Geometrics (SPIG): A Novel Approach Using Heatmaps and Convolutional Neural Network 28
Du Chen, Mariluz Penalva Oscoz, Yang Hai, Martin Rebe Ander, Frans Coenen and Anh Nguyen

Towards Transparent AI in Medical Imaging: Fracture Detection in Hand Radiographs with Grad-CAM Insights 39
Mustafa Juzer Fatehi, Siddharath Malavalli Nagesh, Mandalam Akshit Rao, Stellin John George, J. Angel Arul Jothi and Elakkiya Rajasekar

Redefining Prerequisites Through Text Embeddings: Identifying Practical Course Dependencies 49
Şükrü Kaan Tetik, Emirhan Toprak, Senem Kumova Metin and Hande Aka Uymaz

A Two-Layer Deep Learning Approach for R&D Partner Recommendation in the Self-Driving Vehicle Industry 60
Juite Wang and Ying-Pei Kao

Unsupervised Thematic Context Discovery for Explainable AI in Fact Verification: Advancing the CARAG Framework 67
Manju Vallayil, Parma Nand, Wei Qi Yan and Héctor Allende-Cid

A Convexity-Dependent Two-Phase Training Algorithm for Deep Neural Networks 78
Tomas Hrycej, Bernhard Bermeitinger, Massimo Pavone, Götz-Henrik Wiegand and Siegfried Handschuh

Enhanced LLM Text Classification Method with Embedded Semantic Feature Encoding 87
Meng Wang, Jing Xie, Yang Li, Zhixiong Zhang and Hanyu Li

Hierarchical Patch Compression for ColPali: Efficient Multi-Vector Document Retrieval with Dynamic Pruning and Quantization 98
Bach Duong and Pham Nhat Minh

Retrieval-Augmented Generation in Industry: An Interview Study on Use Cases, Requirements, Challenges, and Evaluation 110
Lorenz Brehme, Benedikt Dornauer, Thomas Ströhle, Maximilian Ehrhart and Ruth Breu

FIRESPARQL: A LLM-Based Framework for SPARQL Query Generation over Scholarly Knowledge Graphs 123
Xueli Pan, Victor de Boer and Jacco van Ossenbruggen

Embeddings Might Be all You Need: Domain-Specific Sentence Encoders for Latin American E-Commerce Questions 135
Rodrigo Caus, Victor Sotelo, Victor Hochgreb and Julio Cesar dos Reis

Suggesting Product Prices in Automotive E-Commerce: A Study Assessing Regression Models and Explicability 147
André Gomes Regino, Gilson Yuuji Shimizu, Fernando Rezende Zagatti, Filipe Loyola Lopes, Rodrigo Bonacin, Julio Cesar Dos Reis and Cristina Dutra de Aguiar

LLMs and Knowledge Discovery in Low-Resource Language Parliamentary Corpora: The PQ Dashboard Case Study 159
Joel Azzopardi

SHORT PAPERS

Experimental Study of Algorithms for Transforming Decision Rule Systems into Decision Trees 173
Kerven Durdymyradov and Mikhail Moshkov

Real-Time Arabic Sign Language Recognition Using YOLOv5 181
Zainab Abualhassan, Haidar Ramadhan, Mohammed Faisal Naji and Hajar Alsulaili

A Hybrid Approach for Mining the Organizational Structure from University Websites 188
Arman Arzani, Theodor Josef Vogl, Marcus Handte and Pedro José Marrón

Triples, Chains, and Trees: A Hybrid Knowledge Representation Framework for Multi-Modal Scientific Content 200
Maodi Hu, Donghuan Song, Xiumin Liu, Xi Sun and Zhixiong Zhang

On the Asymmetrical Nature of Entity Matching Using Pre-Trained Transformers 208
Andrei Olar

Mapping Weaponised Victimhood: A Machine Learning Approach 216
Samantha Butcher and Beatriz De La Iglesia

Assessing Grade Levels of Texts via Local Search over Fine-Tuned LLMs 224
Changfeng Yu and Jie Wang

Predicting Stock Price Movement with LLM-Enhanced Tweet Emotion Analysis 232
An Vuong and Susan Gauch

Objective-Oriented Transformer for Abstractive Document Summarization 240
Parma Nand, CangeGe Zhang and Manju Vallayil

Peer-to-Peer Federated Learning with Trusted Data Sharing for Non-IID Mitigation 248
Mahran Jazi and Irad Ben-Gal

AutoVU-KG: Automated Validation and Updates for Knowledge Graphs with Web-Search-Augmented LLMs 257
Amel Gader and Alsayed Algergawy

Explaining the Judges' Decisions Criteria
Aerty Santos, Gabriel Silveira de Queirós Campos, Cristine Griffo, Eliana Zandonade and Elias de Oliveira 266

CGNTM: Unsupervised Causal Topic Modeling with LLMs and Nonlinear Causal GNNs
Peixuan Men, Longchao Wang, Aihua Li and Xiaoli Tang 275

WPCM: A Multi-Label Patent Classification Method Based on Weakly Supervised Learning
Dechao Wang, Yongjie Li, Jian Zhu and Xiaoli Tang 286

Research on Frontier Discovery of Technological Innovation Based on Knowledge Flow
Dechao Wang, Yongjie Li, Jian Zhu and Xiaoli Tang 294

Weakly Supervised Graph Neural Networks for Scalable 3D Phase Segmentation in Molecular Dynamics Simulations
Abin Shakya and Bijaya B. Karki 302

MultiFlags: A Probabilistic Framework for Article-Based Size Advice in Fashion E-Commerce
Matthias Späth, Andrea Nestler, Henry Böddeker, Leonidas Lefakis, Yevgeniy Puzikov, Rodrigo Weffer, Nour Karessli, Nadja Klein and Reza Shirvany 313

JURISMIND: Context-Driven Retrieval for Accurate and Relevant Legal Question-Answering in Patent Filings
Pandey Shourya Prasad, Vidhish Trivedi, Madhav Rao and Srijoni Sen 323

Contrastive Learning for Conversational Emotion Recognition Using Knowledge Enhancement of Large Language Models
Andrew L. Mackey, B. Israel Cuevas and Susan Gauch 330

From Observations to Causations: A GNN-Based Probabilistic Prediction Framework for Causal Discovery
Rezaur Rashid and Gabriel Terejanu 337

Exploring Customer Service Agent Preferences for Conversational and Keyword-Based Information Retrieval
Nektarios Machner, Yaren Mändle and Florian Matthes 349

JARGES: Detecting and Decoding Jargon for Enterprise Search
Colin Daly and Lucy Hederman 357

Beyond Parameter Counts: Benchmarking Similar-Sized Large Language Models for Next-Item Recommendation
Kavach Dheer, Peter Corcoran and Josephine Griffith 364

Composer Classification Using a Note Difference Graph
Raymond Conlin and Colm O'Riordan 372

From Detection to Diagnosis: A Layered Hybrid Framework for Anomaly Characterization in Maritime Sensor Streams
Nadeem Iftikhar, Cosmin-Stefan Raita, Aziz Kadem, David Buncek, Matthew Haze Trinh, Yi-Chen Lin, Anders Vestergaard and Gianna Belle 380

Uncertainty Estimation and Calibration of a Few-Shot Transfer Learning Model for Lettuce Phenotyping
Rusith Chamara Hathurusinghe Dewage, Habib Ullah, Muhammad Salman. Siddiqui, Rakibul Islam and Fadi-Al Machot 388

Identifying Innovation Frontiers Based on Prediction of Citation Network Links Between Papers and Patents 398
Li Yongjie, Zhu Jian, Wang Longchao and Tang Xiaoli

Towards Interpretable Fairness-Aware Online Machine Learning Through Local Surrogate Models 406
Farnaz Sadeghi, Capri Paquet and Herna Viktor

The WikiWooW Dataset: Harnessing Semantic Similarity and Clickstream-Data for Serendipitous Hyperlinked-Paths Mining in Wikipedia 415
Cosimo Palma and Bence Molnár

Comparing Chain and Tree-Based Reasoning for Explainable Knowledge Discovery in Contract Analytics Using Large Language Models 426
Antony Seabra, Claudio Cavalcante and Sergio Lifschitz

Bias-Mitigating News Search with BiasRank 436
Tim Menzner and Jochen L. Leidner

Decision Rule-Based Learning of Terrorist Threats 448
Nida Meddouri, Loïc Salmon, David Beserra and Elloh Adja

Towards Early Detection of Mild Cognitive Impairment: Predictive Analytics Using the Oculo-Cognitive Addition Test (OCAT) 457
Gaurav N. Pradhan, Sarah E. Kingsbury, Michael J. Cevette, Jan Stepanek and Richard J. Caselli

Integrating Large Language Models into Automated Machine Learning: A Human-Centric Approach 465
Néstor Miguel-Morante, Iván Rivero, Diego García-Prieto, Rafael Duque, Camilo Palazuelos and Abraham Casas

Identifying Research Hotspots and Research Gaps in Specific Research Area Based on Fine-Grained Information Extraction via Large Language Models 473
Yuling Sun, Xuening Cui, Aning Qin and Jiatang Luo

BrandNERD: An Extensive Brand Dataset and Analysis Pipeline for Named Entity Resolution 481
Nicholas Caporusso, Alina Campan, Ayush Bhandari, Stephen Kroeger and Sarita Gautam

Are all Genders Equal in the Eyes of Algorithms? Analysing Search and Retrieval Algorithms for Algorithmic Gender Fairness 489
Stefanie Urchs, Veronika Thurner, Matthias Aßenmacher, Ludwig Bothmann, Christian Heumann and Stephanie Thiemichen

From Free Text to Upper Gastrointestinal Cancer Diagnosis: Fine-Tuning Language Models on Endoscopy and Histology Narratives 501
Kazhan Misri, Leo Alexandre and Beatriz De La Iglesia

RFG Framework: Retrieval-Feedback-Grounded Multi-Query Expansion 509
Ronaldinho Vega Centeno Olivera, Allan M. de Souza and Julio Cesar dos Reis

Prompt Injection Attacks on Large Language Models: Multi-Model Security Analysis with Categorized Attack Types 517
Selin Şaşal and Özgü Can

AUTHOR INDEX 525