COLING 2025

Workshop on Generative AI and Knowledge Graphs (GenAIK)

Proceedings of the Workshop

Abu Dhabi, UAE
January 19, 2025

©Proceedings of The Generative AI and Knowledge Graph Workshop (GenAIK) @COLING-2025

Copyright the International Committee on Computational Linguistics (ICCL) 2025
These proceedings are licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ISBN 979-8-89176-206-0

Preface

Welcome to GenAIK 2025 – The Generative AI and Knowledge Graph Workshop, held in Abu Dhabi, UAE on January 19, 2025.

Generative Artificial Intelligence (GenAI) is a branch of artificial intelligence capable of creating seemingly new, meaningful content, including text, images, and audio. It utilizes deep learning models, such as Large Language Models (LLMs), to recognize and replicate data patterns, enabling the generation of human-like content. Notable families of LLMs include GPT (GPT-3.5, GPT-3.5 Turbo, and GPT-4), LLaMA (LLaMA and LLaMA-2), and Mistral (Mistral and Mixtral). GPT, which stands for Generative Pretrained Transformer, is especially popular for text generation and is widely used in applications like ChatGPT. GenAI has taken the world by storm and revolutionized various industries, including healthcare, finance, and entertainment. However, GenAI models have several limitations, including biases from training data, generating factually incorrect information, and difficulty in understanding complex content. Additionally, their performance can vary based on domain specificity.

In recent times, Knowledge Graphs (KGs) have attracted considerable attention for their ability to represent structured and interconnected information, and have been adopted by many companies in various domains. KGs represent knowledge by depicting relationships between entities, known as facts, usually based on formal ontological models. Consequently, they enable accuracy, decisiveness, interpretability, domain-specific knowledge, and evolving knowledge in various AI applications. The intersection between GenAI and KG has ignited significant interest and innovation in Natural Language Processing (NLP). For instance, by integrating LLMs with KGs during pre-training and inference, external knowledge can be incorporated to enhance the model's capabilities and improve interpretability. When integrated, they offer a robust approach to problem-solving in diverse areas such as information enrichment, representation learning, conversational AI, cross-domain AI transfer, bias, content generation, and semantic understanding. This workshop aims to reinforce the relationships between Deep Learning, Knowledge Graphs, and NLP communities and foster interdisciplinary research in GenAI.

We invited three types of papers: full research papers, short research papers, and position papers. Overall, we received 31 abstract submissions, which were reviewed by 31 members of the Programme Committee. The review process was double-blind. Each paper received three reviews. In total, 15 papers were accepted for publication in this volume, including 12 research papers, 2 short research papers, and 1 position paper.

We would like to express our gratitude to the Organizing Committee and the Program Committee. We would also like to thank our keynote speakers (Kang Liu and Wenya Wang) for accepting our invitations without hesitation and bringing their insights into the importance of knowledge graphs in the age of GenAI. Finally, our gratitude goes also to the sponsor of the conference, NFDI4DataScience, and to the COLING organization team for making the event successful.

January 2025

Genet Asefa Gesese
Harald Sack
Heiko Paulheim
Albert Meroño-Peñuela
Lihu Chen

Organizing Committee

Genet Asefa Gesese, FIZ Karlsruhe, KIT, Germany
Harald Sack, FIZ Karlsruhe, KIT, Germany
Heiko Paulheim, University of Mannheim, Germany
Albert Meroño-Peñuela, King's College London, UK
Lihu Chen, Imperial College London, UK

Program Committee

Paul Groth, University of Amsterdam, The Netherlands
Achim Rettinger, University of Trier, Germany
Cassia Trojahn, Institut de Recherche en Informatique de Toulouse, France
Gerard de Melo, Hasso Plattner Institute, Germany
Pushkar Singh, Google Search Ads AI, United States
Davide Buscaldi, Université Paris 13, France
Pierre-Henri Paris, Paris-Saclay University, France
Mayank Kejriwal, University of Southern California, United States
Finn Arup Nielsen, Technical University of Denmark, Denmark
Marco Bombieri, University of Verona, Italy
Fabien Gandon, Institut national de recherche en informatique et en automatique, France
Giuseppe Rizzo, LINKS Foundation, Italy
Edlira Vakaj Kalemi, Birmingham City University, UK
Vojtech Svatek, Prague University of Economics and Business, Czech Republic
Ondřej Zamazal, Prague University of Economics and Business, Czech Republic
Daniel Garijo, Universidad Politécnica de Madrid, Spain
Sanju Tiwari, TIB Hannover, Germany
Shufan Jiang, FIZ-Karlsruhe, Germany
Jesualdo Fernández-Breis, University of Murcia, Spain
Mauro Dragoni, Fondazione Bruno Kessler, Italy
Kristiina Jokinen, University of Helsinki, Finland
Graham Wilcock, University of Helsinki, Finland
Mathieu d'Aquin, Université de Lorraine, France
Giancarlo Guizzardi, University of Twente, EEMCS, The Netherlands
Oscar Rodríguez Rocha, Institut national de recherche en informatique et en automatique, France
Suvodeep Mazumdar, University of Sheffield, UK
Hiba Arnaout, TU Darmstadt, Germany
Nico Potyka, Cardiff University, UK
Zichao Li, Canoakbit Alliance/University of Waterloo, Canada
Arunav Das, King's College London, UK
Federico Torrielli, University of Turin, Italy

Table of Contents

Effective Modeling of Generative Framework for Document-level Relational Triple Extraction
Pratik Saini and Tapas Nayak 1

Learn Together: Joint Multitask Finetuning of Pretrained KG-enhanced LLM for Downstream Tasks
Anastasia Martynova, Vladislav Tishin and Natalia Semenova 13

GNET-QG: Graph Network for Multi-hop Question Generation
Samin Jamshidi and Yllias Chali 20

SKETCH: Structured Knowledge Enhanced Text Comprehension for Holistic Retrieval
Aakash Mahalingam, Vinesh Kumar Gande, Aman Chadha, Vinija Jain and Divya Chaudhary . . 27

On Reducing Factual Hallucinations in Graph-to-Text Generation Using Large Language Models
Dmitrii Iarosh, Alexander Panchenko and Mikhail Salnikov 43

GraphRAG: Leveraging Graph-Based Efficiency to Minimize Hallucinations in LLM-Driven RAG for Finance Data
Mariam BARRY, Gaetan CAILLAUT, Pierre HALFTERMEYER, Raheel QADER, Mehdi MOUAYAD, Fabrice LE DEIT, Dimitri CARIOLARO and Joseph GESNOUIN 54

Structured Knowledge meets GenAI: A Framework for Logic-Driven Language Models
Farida Helmy Eldessouky, Nourhan Ehab, Carolin Schindler, Mervat Abuelkheir and Wolfgang Minker 66

Performance and Limitations of Fine-Tuned LLMs in SPARQL Query Generation
Thamer Mecharnia and Mathieu d'Aquin 69

Refining Noisy Knowledge Graph with Large Language Models
Na Dong, Natthawut Kertkeidkachorn, Xin Liu and Kiyoaki Shirai 78

Can LLMs be Knowledge Graph Curators for Validating Triple Insertions?
André Gomes Regino and Julio Cesar dos Reis 87

Text2Cypher: Bridging Natural Language and Graph Databases
Makbule Gulcin Ozsoy, Leila Messallem, Jon Besga and Gianandrea Minneci 100

KGFakeNet: A Knowledge Graph-Enhanced Model for Fake News Detection
Anuj Kumar, Pardeep Kumar, Abhishek Yadav, Satyadev Ahlawat and Yamuna Prasad 109

Style Knowledge Graph: Augmenting Text Style Transfer with Knowledge Graphs
Martina Toshevska, Slobodan Kalajdziski and Sonja Gievska 123

Entity Quality Enhancement in Knowledge Graphs through LLM-based Question Answering
Morteza Kamaladdini Ezzabady and Farah Benamara 136

Multilingual Skill Extraction for Job Vacancy–Job Seeker Matching in Knowledge Graphs
Hamit Kavas, Marc Serra-Vidal and Leo Wanner 146

Workshop Program

8:45–8:55 ***Welcome and Opening***

8:55–10:30 **Session 1**

8:55–9:35 ***Keynote 1: Could We Locate Knowledge in Large Language Models?***

09:35–09:55 *Effective Modeling of Generative Framework for Document-level Relational Triple Extraction*
Pratik Saini and Tapas Nayak

9:55–10:10 *Learn Together: Joint Multitask Finetuning of Pretrained KG-enhanced LLM for Downstream Tasks*
Anastasia Martynova, Vladislav Tishin and Natalia Semenova

10:10–10:25 *GNET-QG: Graph Network for Multi-hop Question Generation*
Samin Jamshidi and Yllias Chali

11:00–12:45 **Session 2**

11:00–11:20 *SKETCH: Structured Knowledge Enhanced Text Comprehension for Holistic Retrieval*
Aakash Mahalingam, Vinesh Kumar Gande, Aman Chadha, Vinija Jain and Divya Chaudhary

11:20–11:40 *On Reducing Factual Hallucinations in Graph-to-Text Generation Using Large Language Models*
Dmitrii Iarosh, Alexander Panchenko and Mikhail Salnikov

11:40–12:00 *GraphRAG: Leveraging Graph-Based Efficiency to Minimize Hallucinations in LLM-Driven RAG for Finance Data*
Mariam BARRY, Gaetan CAILLAUT, Pierre HALFTERMEYER, Raheel QADER, Mehdi MOUAYAD, Fabrice LE DEIT, Dimitri CARIOLARO and Joseph GESNOUIN

12:00–12:15 *Structured Knowledge meets GenAI: A Framework for Logic-Driven Language Models*
Farida Helmy Eldessouky, Nourhan Ehab, Carolin Schindler, Mervat Abuelkheir and Wolfgang Minker

12:15–12:35 *Performance and Limitations of Fine-Tuned LLMs in SPARQL Query Generation*
Thamer Mecharnia and Mathieu d'Aquin

13:40–15:30 Session 3

13:40–14:20 ***Keynote 2: Can LLMs function as, enhance and benefit from Knowledge Bases?***

14:20–14:40 *Refining Noisy Knowledge Graph with Large Language Models*
Na Dong, Natthawut Kertkeidkachorn, Xin Liu and Kiyoaki Shirai

14:40–15:00 *Can LLMs be Knowledge Graph Curators for Validating Triple Insertions?*
André Gomes Regino and Julio Cesar dos Reis

15:00–15:20 *Text2Cypher: Bridging Natural Language and Graph Databases*
Makbule Gulcin Ozsoy, Leila Messallem, Jon Besga and Gianandrea Minneci

16:00–17:20 Session 4

16:00–16:20 *KGFakeNet: A Knowledge Graph-Enhanced Model for Fake News Detection*
Anuj Kumar, Pardeep Kumar, Abhishek Yadav, Satyadev Ahlawat and Yamuna Prasad

16:20–16:40 *Style Knowledge Graph: Augmenting Text Style Transfer with Knowledge Graphs*
Martina Toshevska, Slobodan Kalajdziski and Sonja Gievska

16:40–17:00 *Entity Quality Enhancement in Knowledge Graphs through LLM-based Question Answering*
Morteza Kamaladdini Ezzabady and Farah Benamara

17:00–17:20 *Multilingual Skill Extraction for Job Vacancy–Job Seeker Matching in Knowledge Graphs*
Hamit Kavas, Marc Serra-Vidal and Leo Wanner

17:20–17:40 ***Closing***