Knowledge Graphs for Responsible AI

Edlira Vakaj*
Birmingham City University
Birmingham, UK
edlira.vakaj@bcu.ac.uk

Nandana Mihindukulasooriya*
IBM Research
USA, USA
nandana@ibm.com

Manas Gaur
University of Maryland Baltimore County
Baltimore, Maryland, USA
manas@umbc.edu

Arijit Khan
Aalborg University
Aalborg, Denmark
arijitk@cs.aau.dk

Abstract

Responsible AI is built upon a set of principles that prioritize fairness, transparency, accountability, and inclusivity in AI development and deployment. As AI systems become increasingly sophisticated, including the explosion of generative AI, there is a growing need to address ethical considerations and potential societal impacts of their uses. Knowledge graphs (KGs), as structured representations of information, can enhance generative AI performance by providing context, explaining outputs, and reducing biases, thereby offering a powerful framework to address the challenges of responsible AI. By leveraging semantic relationships and contextual understanding, KGs facilitate transparent decision-making, enabling stakeholders to trace and interpret the reasoning behind AI-driven outcomes. Moreover, they provide a means to capture and manage diverse knowledge sources, supporting the development of fair and unbiased AI models. The workshop aims to investigate the role of knowledge graphs in promoting responsible AI principles and creating a cooperative space for researchers, practitioners, and policymakers to exchange insights and enhance their comprehension of KGs' impact on achieving responsible AI solutions. It seeks to facilitate collaboration and idea-sharing to advance the understanding of how KGs can contribute to responsible AI.

Workshop Contact Person. Edlira Vakaj, Birmingham City University, UK. `edlira.vakaj@bcu.ac.uk`

CCS Concepts

• **Information systems** → **Collaborative and social computing systems and tools**; *Graph-based database models*; • **Computing methodologies** → **Knowledge representation and reasoning**.

Keywords

Responsible AI, Interpretability, Privacy, Fairness, Bias Mitigation, Ethical AI, Knowledge Graphs, Large Language Models

---

*Both authors contributed equally to this workshop.

1 Workshop Theme and Topics

We invite submissions of original research, case studies, and position papers on topics related to knowledge graphs and their applications in advancing responsible AI. The workshop explores the intersection of knowledge graphs and ethical considerations in AI development. Submissions may include, but are not limited to, the following topics.

Knowledge Graphs for Bias Mitigation: (1) Techniques and methodologies for using knowledge graphs to identify and mitigate biases in AI models. (2) Case studies demonstrating the successful application of knowledge graphs in addressing bias challenges.

Interpretability and Explainability: (1) Approaches to enhance the interpretability and explainability of black-box AI models through integrating knowledge graphs. (2) Evaluating the effectiveness of knowledge graphs in transparent AI decision-making processes.

Privacy-Preserving Knowledge Graphs: (1) Methods for constructing knowledge graphs that prioritize privacy and comply with data protection regulations. (2) Applications of knowledge graphs in privacy-preserving AI systems.

Fairness in AI with Knowledge Graphs: (1) How knowledge graphs contribute to ensuring fairness in AI applications. (2) Techniques for using KGs and their embeddings to identify and rectify biases in AI models.

Ethical Considerations in Knowledge Graph Construction: (1) Ethical challenges in the creation and maintenance of knowledge graphs. (2) Best practices for ensuring responsible and ethical knowledge graph development. (3) Real-world applications of knowledge graphs in responsible AI.

Integration of Large Language Models (LLMs) and Knowledge Graphs (KGs): (1) Enhancing LLMs' accuracy, consistency, reducing hallucinations and harmful contents generation, fake news detection, fact checking, etc., with knowledge-grounded techniques. (2) Enhancing the interoperability of KG downstream tasks through LLMs' natural language interfaces, transferability, and generalization capacity.

2 Workshop Objectives, Goals, and Expected Outcome

Knowledge graphs (KGs) have been identified as key enablers for explainability in AI [6, 7]. Knowledge graphs and their byproducts, such as KG embeddings, can have their own implicit biases that have to be taken into account [1, 3]. Fot instance, the KG-BIAS workshop [8] at the Automated Knowledge Base Construction (AKBC) conference focuses on identifying biases in automatic KG construction. A deeper integration of knowledge graphs into AI is seen through neuro-symbolic AI. This innovative approach merges the power of statistical machine learning (black-box neural network technologies, e.g., deep learning and LLMs), known for its data-driven predictions, with structured symbolic systems such as knowledge graphs [9, 10]. The Knowledge-infused Learning workshop and the Neuro-Symbolic AI workshop have been central to building a community for transparent and trustworthy AI [5]. Our workshop takes this vision a step further by inviting researchers to explore how to use knowledge graphs to detect hallucinations, falsehoods, contradictions, and knowledge gaps in LLM outputs and contribute to responsible AI principles, e.g., user-in-the-loop, fairness, bias, consistency, and explainability [2, 4, 11].

3 Workshop Length

The proposed length of the workshop is half-day or full-day.

4 Target Audience

The workshop is intended for researchers and practitioners in the broad area of knowledge graphs, semantic web, graph machine learning, data science, deep learning, explainable AI, generative AI, large language models, and responsible AI. We expect around 50 participants in the workshop.

5 Workshop Relevance

Knowledge graphs are fundamental to organizing and retrieving structured information, and can be used to enhance the transparency, fairness, and accountability of AI systems, which are critical issues in the realm of information management, aligning closely with the core themes of CIKM.

6 Past Workshops

- Four editions of the Knowledge Infused Learning Workshop: https://kil-workshop.github.io/
- Three editions of the NLP4KGc Workshop: https://dl.acm.org/doi/fullHtml/10.1145/3543873.3589746

7 Workshop Program Format

The workshop will be held in hybrid mode. Our workshop will have invited (and thus well-prepared) presentations both from academia and industry. They will summarize advances made in the field and lay out some provocative issues related to the area. We will attempt to get speakers from industry (e.g., Google, Meta, Microsoft) who can highlight the issues that industry is facing. The other invited speakers will be leading academics in the area. Bringing industry and academia together is one of the key objectives of the workshop. We hope to augment the flow of ideas between these groups of researchers. We plan to publish the workshop proceedings with all accepted papers in the CEUR proceedings.

8 Program Committee

- Muhammad Afzal, Birmingham City University, UK
- Sule Yildirim Yayilgan, Norwgian University of Science and Teechnology, Norway
- Jagdev Bhogal, Birmingham City University, UK
- Nandana Mihindukulasooriya, IBM Research, USA
- Fernando Ortiz-Rodríguez, Universidad Autonoma de Tamaulipas, Mexico
- Koninika Pal, IIT Palakka, India
- Jeff Z. Pan, University of Edinburgh, UK
- Ali Mohammadi, University of Maryland Baltimore County, USA
- Kaushik Roy, AI Institute, University of South Carolina, USA
- He Tan, Jonkoping University, Sweden
- Sanju Tiwari, Universidad Autonoma de Tamaulipas, Mexico
- Deepa Tilwani, AI Institute, University of South Carolina, USA
- Ramana Malladi, Meta, USA
- Alban Zammit, Paypal, USA
- Naheed Anjum Arafat, National University of Singapore, Singapore
- Bishwamittra Ghosh, Max Planck Institute for Software Systems, Germany
- Simon Razniewski, Max Planck Institute for Informatics, Germany
- Mohamed Abomhara, Norwgian University of Science and Teechnology, Norway
- Chuangtao Ma, Aalborg University, Denmark
- H. V. Jagadish, University of Michigan, USA
- Idir Benouaret, Ecole Pour l'Informatique et les Techniques Avancees, France

More program committee members will be confirmed later.

9 Participation and Selection Process

Each paper will be peer-reviewed by at least three members of the program committee. Then, one co-chair will try to reach consensus among the PC members if there is divergence of opinions. Importance will be given to the impact the work may have if successful.

10 Organizers' Background

Edlira Vakaj is leading the Natural Language Processing AI Lab and is an Associate Professor of Neuro-Symbolic AI at Birmingham City University, UK. She conducts research in multidisciplinary projects focusing on semantic web, knowledge graphs, AI, and semantic data spaces. Edlira is the principal investigator of the ACCORD Horizon project and is engaged in several European and UK-funded projects of various domains where Semantic Web Technologies are applied, such as Renewable Energy (FP7 RENESENG), Industrialised Construction and Industry 4.0 (Innovate UK DfMA Knowledge Transfer Partnership), Higher Education and Youth (Erasmus + Capacity Building, Learning mobility of individuals, Cooperation for innovation and the exchange of good practices action). Dr. Vakaj is an active member of various communities such as The

Alan Turing Knowledge Graph Community, Linked Building Data, Knowledge Graph Creation and Common Action. She organized the 1st NLP4KGC workshop in conjunction with the WWW 2023 conference and the 2nd NLP4KGc workshop in the SEMANTICs 2023 conference.

Nandana Mihindukulasooriya is a senior research scientist at IBM Research, New York, USA. Prior to that, he was a postdoctoral fellow at MIT-IBM Watson AI Lab. He holds a PhD in artificial intelligence from Universidad Politecnica De Madrid, Spain. His research interests include relation extraction and linking, information extraction, knowledge representation and reasoning, and neuro-symbolic AI. Dr. Mihindukulasooriya is also an inventor of several patents related to the same area. He has won several prestigious international awards for the impact of his research contributions to IBM. Nandana was co-organizer of several workshops including the International Workshop on Knowledge Graph Generation from Text at ESWC 2022, International Workshop on Knowledge Graph Summarization at ISWC 2022, SMART Semantic Web Challenge at ISWC 2020-2022, International Workshop on Knowledge Graphs on Travel and Tourism at IWCE 2018, International Workshop on Quality Engineering Meets Knowledge Graph at K-CAP 2017. As a PC member, he has contributed to more than 35 international conferences, including AAAI, IJCAI, ISWC, ESWC, ACL, EMNLP, K-CAP, SEMANTiCS, and others.

Manas Gaur is an assistant professor at the University of Maryland Baltimore County (UMBC) in the Department of Computer Science and Electrical Engineering. He directs the Knowledge-infused AI and Inference (KAI2) lab. Prior to his academic pursuits, Dr. Gaur served as the lead research scientist in Natural Language Processing at the AI Centre in Samsung Research America and he also contributed as a visiting researcher at the prestigious Alan Turing Institute. Dr. Gaur is acknowledged with AAAI New Faculty for 2023 and IEEE Intelligent Systems and Internet Computing best paper awards. He is serving as Guest editor for the IEEE Special Issue on Knowledge-infused Learning and ACM Health Special Issue on Large Language Models, Conversational Systems, and Generative AI in Health.

Arijit Khan is an IEEE senior member, an ACM distinguished speaker, and an associate professor in the Department of Computer Science, Aalborg University, Denmark. He earned his PhD from the Department of Computer Science, University of California, Santa Barbara, USA, and did a post-doc in the Systems group at ETH Zurich, Switzerland. Arijit is the recipient of the prestigious IBM PhD Fellowship in 2012-13, a VLDB Distinguished Reviewer award (2022), and a SIGMOD Distinguished PC award (2024). He published more than 80 papers in premier databases and data mining conferences and journals including ACM SIGMOD, VLDB, IEEE TKDE, IEEE ICDE, SIAM SDM, USENIX ATC, EDBT, The Web Conference (WWW), ACM WSDM, ACM CIKM, ACM TKDD, and ACM SIGMOD Record. He served in the program committee of ACM KDD, ACM SIGMOD, VLDB, IEEE ICDE, IEEE ICDM, EDBT, ACM CIKM, and in the senior program committee of WWW. Dr. Khan served as the co-chair of Big-O(Q) workshop co-located with VLDB 2015 and LLM+KG workshop co-located with VLDB 2024.

Acknowledgments

References

[1] Styliani Bourli and Evaggelia Pitoura. 2020. Bias in knowledge graph embeddings. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. 6–10.

[2] Tingyang Chen, Dazhuo Qiu, Yinghui Wu, Arijit Khan, Xiangyu Ke, and Yunjun Gao. 2024. View-based explanations for graph neural networks. *Proc. ACM Manag. Data* 2, 1 (2024), 40:1–40:27.

[3] Gianluca Demartini. 2019. Implicit bias in crowdsourced knowledge graphs. In *The World Wide Web Conference (WWW)*. 624–630.

[4] Manas Gaur, Keyur Faldu, and Amit P. Sheth. 2021. Semantics of the black-box: Can knowledge graphs help make deep learning systems more interpretable and explainable? *IEEE Internet Computing* 25, 1 (2021), 51–59.

[5] Manas Gaur, Kalpa Gunaratna, Shreyansh P. Bhatt, and Amit P. Sheth. 2022. Knowledge-infused learning: A sweet spot in neuro-symbolic AI. *IEEE Internet Computing* 26, 4 (2022), 5–11.

[6] Freddy Lécué. 2020. On the role of knowledge graphs in explainable AI. *Semantic Web* 11, 1 (2020), 41–51.

[7] Enayat Rajabi and Kobra Etminani. 2022. Knowledge-graph-based explainable AI: A systematic review. *Journal of Information Science* (2022).

[8] Amit P. Sheth, Kaushik Roy, and Manas Gaur. 2023. Neurosymbolic artificial intelligence (why, what, and how). *IEEE Intelligent Systems* 38, 3 (2023), 56–62.

[9] Ilaria Tiddi and Stefan Schlobach. 2022. Knowledge graphs as tools for explainable machine learning: A survey. *Artif. Intell.* 302 (2022), 103627.

[10] Yuxiang Wang, Arijit Khan, Tianxing Wu, Jiahui Jin, and Haijiang Yan. 2020. Semantic guided and response times bounded top-k similarity search over knowledge graphs. In *IEEE International Conference on Data Engineering (ICDE)*. 445–456.

[11] Ruwan Wickramarachchi, Cory A. Henson, and Amit P. Sheth. 2020. An evaluation of knowledge graph embeddings for autonomous driving data: Experience and practice. In *AAAI Symposium on Combining Machine Learning and Knowledge Engineering in Practice (AAAI-MAKE), Volume I (CEUR Workshop Proceedings, Vol. 2600)*.