

Joint Proceedings of the First Workshop on LLM-driven Knowledge Graph and Ontology Engineering (llms4kgoe) and the Third workshop on Evaluation of Language Models in Knowledge Engineering (ELMKE)

Abstract

This volume contains the joint proceedings of the first Workshop on LLM-driven Knowledge Graph and Ontology Engineering (LLMS4KGOE 2026) and the third Workshop on Evaluation of Language Models in Knowledge Engineering (ELMKE 2026), held in conjunction with the 23rd European Semantic Web Conference (ESWC 2026) in Dubrovnik, Croatia. LLMS4KGOE 2026 addresses the emerging intersection of large language models, semantic technologies, knowledge graphs, and ontology engineering, covering LLM-assisted ontology creation, refinement, validation, hallucination mitigation, semantic consistency, provenance, governance, and human-in-the-loop methods. ELMKE 2026 focuses on the evaluation of language models in knowledge engineering tasks, including knowledge extraction, knowledge base construction, curation, completion, and generative tasks, with particular attention to standardization, reproducibility, transparency, human evaluation, and benchmark development. Together, the workshops reflect a growing research agenda around the responsible, rigorous, and practically useful integration of language models into knowledge engineering and Semantic Web workflows. The papers collected in this volume contribute methods, benchmarks, systems, and critical analyses that advance both the construction and evaluation of LLM-supported knowledge graphs and ontologies.

Keywords

Language Models, Knowledge Engineering, Knowledge Graphs, Ontology Engineering, Semantic Web, Evaluation, Benchmarks, Human-in-the-Loop, Hallucination Mitigation, Provenance, LLM-driven Ontology Construction

Preface

This volume contains the joint proceedings of the *First Workshop on LLM-driven Knowledge Graph and Ontology Engineering (LLMS4KGOE 2026)* and the *Third Workshop on Evaluation of Language Models in Knowledge Engineering (ELMKE 2026)*. These workshops were co-located with the 23rd European Semantic Web Conference (ESWC 2026), which took place in Dubrovnik, Croatia, from May 10 to 14, 2026.

First Workshop on LLM-driven Knowledge Graph and Ontology Engineering (LLMS4KGOE 2026). The Workshop on LLM-driven Knowledge Graph and Ontology Engineering focuses on the emerging intersection of Large Language Models (LLMs), semantic technologies, knowledge graphs, and ontology engineering. Traditional ontology engineering is a difficult, time-consuming, and expert-driven process, while the advent of LLMs offers new opportunities to accelerate and democratize ontology development. At the same time, LLM-generated ontologies raise critical challenges related to quality, coherence, formal correctness, evaluation, and trustworthiness. The workshop provides a forum for discussing both theoretical and practical aspects of LLM-based ontology construction, including semi-automated and fully automated ontology generation pipelines, ontology refinement and validation, evaluation methodologies, hallucination mitigation, provenance, governance, human-in-the-loop protocols, and neuro-symbolic approaches.

Third Workshop on Evaluation of Language Models in Knowledge Engineering (ELMKE 2026). The Workshop on Evaluation of Language Models in Knowledge Engineering addresses the growing need for robust, transparent, and reproducible evaluation methods for the use of language models in knowledge engineering tasks. Language models have shown promise in areas such as knowledge

Joint Proceedings of the LLMS4KGOE 2026 and ELMKE 2026 Workshops, co-located with the European Semantic Web Conference (ESWC 2026), May 10-14, 2026, Dubrovnik, Croatia



© 2026 Copyright for the individual papers by its authors. Copyright for the volume as a collection by its editors. This volume and its papers are published under the Creative Commons License Attribution 4.0 International (CC BY 4.0).

extraction, knowledge base construction, curation, completion, and generative tasks. However, evaluating LM-generated results remains challenging due to the lack of comprehensive and formally defined evaluation frameworks, the reliance on human effort, and the difficulty of comparing methods and reproducing experiments. ELMKE aims to support a community-driven effort toward standardizing evaluation practices, developing benchmarks, improving transparency, and fostering collaborative platforms and dashboards for evaluating language models in knowledge engineering.

Together, LLMS4KGOE 2026 and ELMKE 2026 reflect the growing importance of large language models in the Semantic Web and knowledge engineering communities. While LLMS4KGOE emphasizes the construction, refinement, and governance of LLM-supported knowledge graphs and ontologies, ELMKE focuses on the evaluation principles, methods, and benchmarks needed to assess such systems rigorously. The papers included in this volume contribute to a shared research agenda on the responsible, reliable, and practically useful integration of language models into knowledge graph and ontology engineering workflows.

The two workshops received 19 submissions in total. All submitted papers underwent a peer-review process, where each paper was reviewed by experts from the respective program committees. This volume contains the accepted papers, which are distributed across the workshops as follows:

- 11 submissions were received for LLMS4KGOE 2026, of which 9 were accepted: 7 full papers and 2 short papers;
- 8 submissions were received for ELMKE 2026, of which 7 were accepted.

The workshops were held on May 11, 2026, in conjunction with ESWC 2026 in Dubrovnik, Croatia. Each workshop featured paper presentations and discussions on the role of language models in knowledge graph construction, ontology engineering, and the evaluation of language models in knowledge engineering tasks. We sincerely thank all authors, presenters, reviewers, program committee members, organizers, and participants for their contributions, which fostered engaging and insightful conversations across both workshops.

Review Process and Important Dates

LLMS4KGOE 2026 LLMS4KGOE 2026 employed a single-blind peer-review process. Submitted manuscripts were evaluated by members of the Program Committee based on their relevance to the workshop topics, originality, technical quality, clarity of presentation, and potential impact on the fields of knowledge graph and ontology engineering. Each submission received reviews from at least three reviewers.

The important dates for LLMS4KGOE 2026 were as follows:

- Submission deadline: March 16, 2026 (extended)
- Notification of acceptance: April 20, 2026
- Camera-ready deadline: May 5, 2026

ELMKE 2026 ELMKE 2026 employed a peer-review process conducted by members of the Program Committee. Submitted manuscripts were assessed with respect to their relevance to the workshop themes, originality, technical quality, clarity, reproducibility, and potential impact on the evaluation of language models in knowledge engineering. Each submission received reviews from at least three reviewers.

The important dates for ELMKE 2026 were as follows:

- Submission deadline: March 17, 2026
- Notification of acceptance: March 31, 2026
- Camera-ready deadline: April 15, 2026

Editors

Aryan Singh Dalal, Kansas State University, USA

Kathleen Jagodnik, Kansas State University, USA

Maria Maleshkova, Helmut Schmidt University,
Germany

Cogan Shimizu, Wright State University, USA

Anna Sofia Lippolis, University of Bologna, Italy

Reham Alharbi, Taibah University, Saudi Arabia

Bohui Zhang, King's College London, UK

Yuan He, Amazon, USA

Hande Küçük McGinty, Kansas State University,
USA

Thread I: First Workshop on LLM-driven Knowledge Graph and Ontology Engineering (LLMS4KG OE 2026)

Workshop Organisation of LLMS4KG OE 2026

General Chairs

- Aryan Singh Dalal, Kansas State University, Manhattan, Kansas, US
- Cogan Shimizu, Wright State University, US
- Maria Maleshkova, Helmut Schmidt University / Universität der Bundeswehr Hamburg, DE
- Kathleen Jagodnik, Kansas State University, US
- Hande McGinty, Kansas State University, Manhattan KS, US

Program Committee

- Antrea Christou, Wright State University, Ohio, US
- Miguel Ceriani, University of Bari, IT
- Robert Buchmann, Babeş-Bolyai University of Cluj Napoca, RO
- Vaibhav Gupta, Helmut Schmidt Universität – Universität der Bundeswehr Hamburg, DE
- Anna Sofia Lippolis, CNR ISTC, IT
- Yinglun Zhang, Kansas State University, US
- Alessandro Oltramari, Bosch Research and Technology Center, US
- Chris Davis Jaldi, Wright State University, US
- Pravesh Niroula, Kansas State University, US
- Anmol Saini, Wright State University, US
- Srikar Reddy Gadusu, Kansas State University, US
- María Poveda-Villalón, Universidad Politécnica de Madrid, ES
- Nirmal Gelal, Kansas State University, US
- Daniel Garijo, Universidad Politécnica de Madrid, ES
- Simon Burbach, Helmut Schmidt University / University of the Federal Armed Forces Hamburg, DE
- Paul Groth, University of Amsterdam, NL
- Maaike de Boer, TNO, NL
- Nicole Obretincheva, King's College London, UK
- Soheil Abadifard, Kansas State University, US
- Christian Bizer, University of Mannheim, DE
- Yihang Zhao, King's College London, UK
- Alexis Ellis, Wright State University, US
- Hanna Abi Akl, Université Côte d'Azur, Inria, CNRS, I3S / Data ScienceTech Institute, FR
- Aldo Gangemi, Università di Bologna & CNR-ISTC, IT

Keynote

Presentations

- *Ground-truth Construction and Evaluation of LLM Contribution to Life Sciences Tool Annotation*, by Ulysse Le Clanche, Melvin Selim Atay, Elise Bannier, Anaïs Baudot, Lea Bellenger, Alexandrina Bodrug, Samuel Chaffron, Eric Charpentier, Erwan Corre, Clémence Frioux, Aurélie Lardenois, Frédéric Lemoine, Camille Maumet, Cyril Noël, Perrine Paul-Gilloteaux, Paul Simion, Morgane Térézol, Olivier Dameron, and Alban Gaignard. This work constructs an expert-validated ground-truth dataset for life sciences software annotation and evaluates how LLM-generated annotations complement domain-expert contributions.

- *Benchmarking Resource-Efficient LLMs for Research Topic Ontology Generation in the Biomedical Field*, by Tanay Aggarwal, Angelo Salatino, Francesco Osborne, and Enrico Motta. This paper evaluates small open-source LLMs for identifying semantic relations between biomedical research topics and shows that targeted fine-tuning substantially improves ontology-generation performance.
- *Automating AI Risk Profiling with LLM-Engineered OWL 2 Axioms*, by Manas Goyal, Yogesh Kamath, and Swathi Shyam Sunder. This short paper presents a neuro-symbolic workflow in which LLM-generated OWL axioms and DL reasoning automate industrial AI risk profiling from questionnaire-based business rules.
- *Pitfalls in AI-Generated Ontologies: Strategies for Detection and Mitigation*, by Pasquale Lisena, Julien Plu, Oscar Moreno Escobar, Edouard Trouilleux, and Raphaël Troncy. This work identifies ontology pitfalls specific to LLM-generated ontologies and proposes detection strategies combining structural analysis, lexical similarity, and LLM-based evaluation.
- *PERSEUS: Reducing LLM Hallucinations in Multimodal Knowledge Graph Construction and Ontology Learning*, by Aryan Singh Dalal and Hande Kucuk McGinty. This short paper introduces a multimodal knowledge graph construction framework that reduces hallucinated triples through hybrid retrieval and NLI-based source-evidence verification.
- *OG-NSD: Neuro-Symbolic Ontology Drafting from Natural-Language Requirements*, by Efstratios Skaperdas, Kristi Cami, and Nick Bassiliades. This paper presents a neuro-symbolic pipeline that combines LLM-based ontology drafting with SHACL validation, OWL reasoning, competency-question testing, and iterative symbolic repair.
- *MASEO: A Multi-Agent System for Explainable Ontology Generation*, by Jiayi Li, Ziyuan Wang, Daniel Garijo, and María Poveda Villalón. This work proposes a multi-agent ontology generation framework that transforms competency questions into OWL ontologies while tracking provenance for each generated term and axiom.
- *IDEA2: Expert-in-the-loop Competency Question Elicitation for Collaborative Ontology Engineering*, by Elliott Watkiss-Leek, Reham Alharbi, Harry Rostron, Andrew Ng, Ewan Johnson, Andrew Mitchell, Terry R. Payne, Valentina Tamma, and Jacopo de Berardinis. This paper introduces an expert-in-the-loop workflow that uses LLMs and collaborative feedback to iteratively elicit, review, and refine competency questions for ontology engineering.
- *Towards Automated Ontology Generation from Unstructured Text: A Multi-Agent LLM Approach*, by Abid Talukder, Maruf Ahmed Mridul, and Oshani Seneviratne. This work studies automated ontology generation from unstructured insurance contracts and shows that a planning-first multi-agent LLM architecture improves structural ontology quality.

Acknowledgments

We would like to sincerely thank all the authors for their valuable contributions and the members of the Program Committee for their diligent work in the review process.

Declaration on Generative AI

During the preparation of this preface, the editors may have used AI-powered tools for assistance with language and formatting. After using these tools, the editors reviewed and edited the content as needed and take full responsibility for the content of this publication.

Thread II: Third Workshop on Evaluation of Language Models in Knowledge Engineering (ELMKE 2026)

Workshop Organisation of ELMKE 2026

The Third Workshop on Evaluation of Language Models in Knowledge Engineering (ELMKE 2026) was co-located with the 23rd European Semantic Web Conference (ESWC 2026) in Dubrovnik, Croatia. The workshop focused on the evaluation of language models in knowledge engineering tasks, including knowledge extraction, knowledge base construction, curation, completion, and generative tasks. ELMKE 2026 aimed to support a community-driven effort toward standardizing evaluation practices for language models in knowledge engineering, with particular attention to reproducibility, transparency, human evaluation, benchmark development, and collaborative evaluation platforms.

General Chairs

- Anna Sofia Lippolis, University of Bologna, IT
- Reham Alharbi, Taibah University, SA
- Bohui Zhang, King's College London, UK
- Yuan He, Amazon, USA

Program Committee

- Terry Payne, University of Liverpool, UK
- Arunav Das, King's College London, UK
- Daniel Dobriy, Vienna University of Economics and Business, AT
- Hamed Babaei Giglou, TIB Leibniz Information Centre for Science and Technology, DE
- Gerrit Quaremba, King's College London, UK
- Célian Ringwald, Université Côte d'Azur, Inria, CNRS, I3S, FR
- George Hannah, University of Liverpool, UK
- Desiree Heim, DFKI, DE
- Mohammad Javad Saeedizade, Linköping University, SE
- Yihang Zhao, King's College London, UK
- Arianna Graciotti, University of Groningen, NL

Keynote: Designing Better Knowledge Engineering Assistants

Knowledge engineering focuses on the creation and stewardship of knowledge-based systems. As a field, it occupies a unique niche between software engineering, which involves crafting software that represents knowledge computationally, and AI, where software can reason upon knowledge representations to emulate human thought. Like many areas of knowledge work, the field is being reshaped by generative AI. This brings new opportunities to address long-standing challenges of scale and inclusivity, while also raising concerns about accuracy, bias, and the responsible use of automation. In this talk, I will explore recent work in AI and human computer interaction that examines how to design better knowledge engineering assistants.

Keynote speaker Elena Simperl is a Professor of Computer at King's College London and the Director of Research for the Open Data Institute (ODI). She is a Fellow of the British Computer Society and the Royal Society of Arts, and a Hans Fischer Senior Fellow. Elena's work is at the intersection between AI and social computing. She features in the top 100 most influential scholars in knowledge engineering of the last decade and in the Women in AI 2000 ranking. She is the president of the Semantic Web Sciences Association.

Presentations

- *Understanding the Gap in Evaluation Frameworks and Metrics for RAG-based Question-Answering System* by Nakul Mehta, Adegboyega Ojo and Edward Curry
- *T2S-Metrics: Unified Library for Evaluating SPARQL Queries Generated From Natural Language* by Yousouf Taghzouti, Tao Jiang, Camille Juigné, Benjamin Navet, Fabien Gandon, Franck Michel and Louis Felix Nothias
- *A Benchmark and Field Study on the Use of LLMs and RAG for Harmonizing Job Descriptions in the Austrian Federal Administration* by Nicolas Ferranti, Andreas Mild and Stefan Sobernig
- *Ontology Reuse in LLMs Generated Ontologies: A Comparative Evaluation Framework* by Henon Mengistu Lamboro, Amel Bouzeghoub, Cécile Rabrait, Antonio Kung, Olivier Genest and Amélie Gyrard
- *Towards an Evaluation Framework for Generated Formal Knowledge* by Maxime Haurel, Armelle Brun and Mathieu D'Aquin
- *Evaluating LLMs in Ontology Transformation: Study on Property Chain Shortcutting* by Kateřina Haniková, Vojtěch Svátek and Ondřej Zamazal
- *LLM Driven Justified Alignment* by Diego Conde Herreros, George Hannah, Terry Payne, Jacopo de Berardinis, Valentina Tamma, David Chaves-Fraga and Oscar Corcho

Acknowledgments

We would like to thank the authors for their valuable contributions, the Program Committee for their rigorous reviews, and all participants for making the third ELMKE workshop an engaging and successful event.

Declaration on Generative AI

During the preparation of this preface, the editors may have used AI-powered tools for assistance with language and formatting. After using these tools, the editors reviewed and edited the content as needed and take full responsibility for the content of this publication.