Preface for the Fifth International Workshop on LLM-integrated Knowledge Graph Generation from Text

Sanju Tiwari[1,3], Nandana Mihindukulasooriya[2], Jennifer D'Souza[3] and Francesco Osborne[4]

[1]*Sharda University, Delhi-NCR, India*
[2]*IBM Research, USA*
[3]*TIB Leibniz Information Centre for Science and Technology, Germany*
[4]*KMi, The Open University & University of Milano-Bicocca, Italy*

TEXT2KG is the Fifth International Workshop on LLM-Integrated Knowledge Graph Generation From Text (TEXT2KG), co-located with the Extended Semantic Web Conference (ESWC 2026). The primary aim of this workshop is to bring together researchers from multiple disciplines, including Natural Language Processing (NLP), Large Language Models (LLMs), Entity Linking (EL), Relation Extraction (RE), Knowledge Representation and Reasoning (KRR), Knowledge Graph Construction (KGC), Semantic Web, Linked Data, Graph Neural Networks (GNNs), Information Extraction, and other related areas, to foster discussion and advance the state of the art in knowledge graph generation from text.

The workshop featured an invited keynote by Raphael Troncy (EURECOM, France) titled *"From Strings to Knowledge Graphs: What Text2KG Still Gets Wrong"*. The organizing team is grateful to everyone who contributed to making TEXT2KG 2026 a success. We sincerely thank the organizers of ESWC 2026, the Program Committee members for their rigorous review process, which ensured an average of three reviews per submission, the keynote speaker, the steering and advisory committee (Amit Sheth, Sören Auer, Enrico Motta, Anna Fensel, Maria Esther Vidal, Edlira Vakaj, Fernando Ortiz-Rodriguez, Sven Groppe, Dimitris Kontokostas, Nandana Mihindukulasooriya, and Mayank Kejriwal), and the Publicity Chair (Joey Yip, Ronak Panchal) for their invaluable support and contributions.

TEXT2KG 2026 received 18 submissions, of which 12 papers were accepted following a rigorous peer-review process. Each submission was reviewed by at least three experts from diverse research backgrounds. The accepted papers presented at the workshop are listed below:

- From Text to Triples: A Neuro-Symbolic Pipeline for Structured Claim Extraction
- Mapping Latent Narrative Content Using LLMs: Essay on Economic Uncertainty Drivers
- CIExMAS: Closed Information Extraction using Multi Agent Systems

ESWC'26: 5th Workshop on LLM-Integrated Knowledge Graph Generation from Text (TEXT2KG), May 10-14, 2026 | Dubrovnik, Croatia
tiwarisanju18@ieee.org (S. Tiwari); nandana.m@ibm.com (N. Mihindukulasooriya); Jennifer.DSouza@tib.eu (J. D'Souza); francesco.osborne@open.ac.uk (F. Osborne)

- A Retrieval-Augmented LLM Pipeline for Generating Cognitive Conceptualization Knowledge Graphs
- xch-MIND: Multi-Agent Interpretive Knowledge Graph Augmentation over Cultural Heritage Linked Data
- Augmenting Datasets for Fine-Tuning Large Language Models Using Semantic Variations
- Leveraging Graph Structure in Seq2Seq Models for Knowledge Graph Link Prediction
- Using LLMs to Develop Legal Rule Ontologies for Compliance Checking
- Integrating LLMs and the PrOntoNarr Ontology for Automated KG Construction in Multi-level Propaganda Narrative Analysis
- An Agentic Framework for Dynamic Ontology Generation
- From "Strings" to "Things" for Personal Knowledge Graphs: Evaluating LLM Triple Extraction for Recommendation Systems
- LLM-Assisted Multilingual Information Extraction for Knowledge Graph Population: The DBpedia Hindi Chapter

Keynote
Raphael Troncy, EURECOM, France

Talk Title "From Strings to Knowledge Graphs: What Text2KG Still Gets Wrong"

Abstract
Knowledge graph construction from text has long been framed as a pipeline of entity recognition, entity linking, and relation extraction. Recent progress in large language models has radically simplified the production of plausible triples, and systems such as Microsoft GraphRAG have popularized the use of automatically generated graphs for retrieval-augmented question answering. Yet the central challenge remains: how do we move from plausible text-derived triples to trustworthy, reusable, ontology-compliant knowledge graphs? In this talk, the Text2KG task is revisited from both Natural Language Processing (NLP) and Semantic Web perspectives. Open information extraction, which supports exploratory graph construction, is contrasted with ontology-guided information extraction, where extracted knowledge must satisfy explicit conceptual models, domain and range constraints, provenance requirements, and graph-level consistency. Neural and generative approaches to relation extraction, including BERT-based models and REBEL-like sequence-to-sequence models, are discussed, highlighting the need for richer extraction paradigms involving causal relations, events, frames, and n-ary structures for constructing high-quality domain knowledge graphs.

The talk presents several case studies, including ontology-based extraction of olfactory heritage knowledge in ODEUROPA, multimodal knowledge graph population in SILKNOW, and recent work on Text2KGBench, its re-annotation, and the development of the PERSEUS model in collaboration with Lettria. These examples demonstrate both the promise and the limitations of current LLM-based approaches, emphasizing challenges related to faithfulness, schema conformance, evaluation, and reproducibility. The talk concludes by arguing that the next generation of Text2KG systems should be evidence-grounded, ontology-aware, multimodal, and benchmark-driven, and introduces the TRIPLET challenge and its joint text–table–knowledge

graph information extraction task as an important step toward evaluating knowledge extraction in heterogeneous real-world documents.

Collocated Events

- Third Biochemical Knowledge Extraction Challenge (BiKE26)
- Second Text2SPARQL Challenge (Text2SPARQL26)

Best Paper Award

- Based on the evaluations of the reviewers and the recommendations of the organizing committee, the Best Paper Award was presented to:
  - **CIExMAS: Closed Information Extraction using Multi Agent Systems**. *Max Lautenbach and Sven Hertling*

Runner-up Paper Awards

- **A Retrieval-Augmented LLM Pipeline for Generating Cognitive Conceptualization Knowledge Graphs**. *Simone Pinna, Silvia Maria Massa, Diego Reforgiato, and Daniele Riboni*
- **xch-MIND: Multi-Agent Interpretive Knowledge Graph Augmentation over Cultural Heritage Linked Data**. *Michele Stingo*

Organizing Committee

- Sanju Tiwari, Sharda University, Delhi-NCR, India & TIB Hannover, Germany
- Nandana Mihindukulasooriya, IBM Research, USA
- Jennifer D'Souza, TIB Leibniz Information Centre for Science and Technology, Germany
- Francesco Osborne, KMi, The Open University & University of Milano-Bicocca, Italy

Steering and Advisory Committee

- Amit Sheth, University of South Carolina, USA
- Sören Auer, Leibniz University of Hannover and TIB, Germany
- Enrico Motta, The Open University, UK
- Anna Fensel, Wageningen University & Research, The Netherlands, & University of Innsbruck, Austria
- Maria Esther Vidal, Leibniz University of Hannover and TIB, Germany
- Edlira Vakaj, Birmigham City University, UK
- Fernando Ortiz-Rodriguez, Universidad Autonoma de Tamaulipas, Mexico
- Sven Groppe, University of Lübeck, Germany
- Dimitris Kontokostas, Medidata, Greece
- Mayank Kejriwal, University of Southern California, USA.

Publicity Chair

- Hong Yung (Joey) Yip, AIISC, University of South Carolina, USA

- Ronak Panchal, Cognizant, India

Program Committee

- Ahlem Rhayem, DLR, Germany
- Anelia Kurteva, University of Birmingham, UK
- Daniil Dobriy, Vienna University of Economics and Business, Vienna, Austria
- Davide Buscaldi, LIPN, Université Paris 13, Sorbonne Paris Cité, Paris
- Dimitris Kontokostas, Medidata, Greece
- Disha Purohit, TIB – Leibniz Information Centre for Science and Technology and University Library, Hannover, Germany
- Edgard Marx, Byondis, Netherlands
- Fatima Zahra Amara, University of Abbes Laghrour Khenchela, Algeria
- Fatma Chamekh, French National Survey, France
- Francesco Osborne, KMi, The Open University, UK
- Hamed Babaei Giglou, TIB Leibniz Information Centre for Science and Technology, Germany
- Hong Yung Yip, Artificial Intelligence Institute, University of South Carolina
- Jennifer D'Souza, TIB Leibniz Information Centre for Science and Technology University Library, Hannover, Germany
- Maosheng Guo, Harbin Institute of Technology, China
- Marlene Goncalves, Universidad Simón Bolívar, Venezuela
- Mauro Dragoni, Fondazione Bruno Kessler
- Mayank Kejriwal, University of Southern California, USA
- Muhammad Raza Naqvi, Universite de Toulouse Ecole Nationale d'Ingénieurs de Tarbes ENIT, Tarbes
- Nandana Mihindukulasooriya, IBM Research AI
- Sven Groppe, University of Luebeck, Germany
- Tek Raj Chhetri, Massachusetts Institute of Technology, USA
- Tommaso Soru, University of Leipzig, Germany

Acknowledgements The editors would like to wholeheartedly thank supporting keynote, advisory team, authors, the program committee and other organizers for their constant support to make this event successful.