

TextGraphs @ ACL 2024

**Proceedings of TextGraphs-17: Graph-based Methods for
Natural Language Processing**

**The 62nd Annual Meeting of the Association of
Computational Linguistics**

August 15, 2024

The TextGraphs @ ACL organizers gratefully acknowledge the support from the following sponsors.

Shared Task



©2024 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-145-2

Preface

Welcome to the seventeenth edition of TextGraphs, the workshop on graph-based methods for natural language processing. This edition was organized on August 15, 2024, in Bangkok, Thailand, co-located with the 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024). **TextGraphs-17 commemorates Prof. Dragomir Radev (1968–2023) and his seminal contributions to the TextGraphs community and the natural language processing field.**

Recent years have witnessed remarkable advances in natural language processing (NLP) and graph theory domains that mostly develop independently with rare intersections. The seventeenth edition of the TextGraphs workshop surpassed its boundaries, widening the workshop topic coverage to capture not only a well-studied graph domain but a more general yet underexplored structured data domain as well. Besides, we strongly encouraged novel research efforts that aim to explore the hot topic of the large language model (LLM) prompting from the unique perspective of graph theory. Thus, our workshop sought to establish more mutually beneficial relationships between NLP and structured data to address the pivotal limitations of each domain.

Toward this end, the key topic for TextGraphs-17 was **Knowledge Graphs meet LLMs**. A proper utilization of graph-based methods for reasoning over a Knowledge Graph (KG) is a prospective way to overcome critical limitations of the existing LLMs which lack interpretability and factual knowledge and are prone to the hallucination problem. Vice versa, the incorporation of LLM knowledge discovered from large textual collections may help many graph-related tasks, such as KG completion and graph representation learning. This topic calls for research on the joint use of KG and LLM for improved processing of either the NLP or graph domain. In line with this topic, a shared task on KGQA was organized and co-located with the workshop.

We highlight three more topics relevant to the TextGraphs workshop:

- **Chain Prompting of LLMs.** Recent studies show that prompting strategies like Chain-of-Thought and Graph-of-Thought enhance language understanding and generation tasks compared to the traditional few-shot methods. This topic calls for research on advanced prompting schemes and software for LLMs and other pre-trained machine learning models.
- **Learning from Structured Data.** This topic calls for research that bridges textual and structured data formats, including relational and non-relational databases, as well as standardized data formats (such as XML, JSON, RDF, etc.)
- **Interpretability of NLP Systems.** This topic calls for research which adopts structured data and employs graph-based methods to shed light on decision-making and logic behind modern LLMs. This includes work on applying a KG to explore and evaluate factual awareness, treating the interpretability problem from the GT perspective, or other applications of graphs to make LLMs more understandable.

Besides the main paper presentations, we ran a shared task on Knowledge Graph Question Answering (KGQA). In this task, given a textual question and a list of entities with the corresponding KG subgraphs, the participating system should choose the entity that correctly answers the question. Our competition attracted thirty teams, four of which outperformed our strong ChatGPT-based zero-shot baseline.

Overall, this year, we received 25 submissions, out of which 8 submissions were accepted, 7 submissions were accepted as shared task papers (including a report by shared task organizers), 8 submissions were rejected by the reviewers, and 2 submissions were desk rejected.

We want to thank our keynote speaker, Rada Mihalcea, and we are also thankful to the members of the program committee for their valuable and high-quality reviews. Their expert feedback has benefited all submissions. Their timely contribution was the basis for accepting an excellent list of papers and making the seventeenth edition of TextGraphs a success. Also, we are grateful to JetBrains, which provided prizes for the competition participants.

Dmitry Ustalov, Yanjun Gao, Alexander Panchenko, Elena Tutubalina, Irina Nikishina, Arti Ramesh, Andrey Sakhovskiy, Ricardo Usbeck, Gerald Penn, Marco Valentino

TextGraphs-17 Organizers

August 2024

Organizing Committee

General Chair

Dmitry Ustalov, JetBrains, Serbia

Program Chairs

Alexander Panchenko, Artificial Intelligence Research Institute, Russia

Elena Tutubalina, Artificial Intelligence Research Institute, Russia

Irina Nikishina, University of Hamburg, Germany

Arti Ramesh, Binghamton University, USA

Andrey Sakhovskiy, Kazan Federal University, Russia

Ricardo Usbeck, University of Hamburg, Germany

Gerald Penn, University of Toronto, Canada

Marco Valentino, Idiap Research Institute, Switzerland

Publication Chair

Yanjun Gao, University of Wisconsin-Madison, USA

Keynote Talk
**A Legacy of Graphs: Celebrating Dragomir Radev's
Contributions to Graph-Based Natural Language Processing**

Rada Mihalcea

University of Michigan

2024-08-15 14:00:00 – Room: Centara Grand and Bangkok Convention Centre, Thailand

Bio: Rada Mihalcea is the Janice M. Jenkins Professor of Computer Science and Engineering at the University of Michigan and the Director of the Michigan Artificial Intelligence Lab. Her research interests are in natural language processing, with a focus on multimodal processing and computational social sciences. She is an ACM Fellow, a AAAI Fellow, and served as ACL President (2018–2022 Vice/Past). She is the recipient of a Sarah Goddard Power award (2019) for her contributions to diversity in science, and the recipient of a Presidential Early Career Award for Scientists and Engineers awarded by President Obama (2009).

Table of Contents

<i>Learning Human Action Representations from Temporal Context in Lifestyle Vlogs</i> Oana Ignat, Santiago Castro, Weiji Li and Rada Mihalcea	1
<i>ConGraT: Self-Supervised Contrastive Pretraining for Joint Graph and Text Embeddings</i> William Brannon, Wonjune Kang, Suyash Fulay, Hang Jiang, Brandon Roy, Deb Roy and Jad Kabbara	19
<i>A Pipeline Approach for Parsing Documents into Uniform Meaning Representation Graphs</i> Jayeol Chun and Nianwen Xue	40
<i>Financial Product Ontology Population with Large Language Models</i> Chanatip Saetia, Jiratha Phruetthiset, Tawunrat Chalothorn, Monchai Lertsutthiwong, Supawat Taerungruang and Pakpoom Buabthong	53
<i>Prompt Me One More Time: A Two-Step Knowledge Extraction Pipeline with Ontology-Based Verification</i> Alla Chepurova, Yuri Kuratov, Aydar Bulatov and Mikhail Burtsev	61
<i>Towards Understanding Attention-based Reasoning through Graph Structures in Medical Codes Classification</i> Noon Pokaratsiri Goldstein, Saadullah Amin and Günter Neumann	78
<i>Leveraging Graph Structures to Detect Hallucinations in Large Language Models</i> Noa Nonkes, Sergei Agaronian, Evangelos Kanoulas and Roxana Petcu	93
<i>Semantic Graphs for Syntactic Simplification: A Revisit from the Age of LLM</i> Peiran Yao, Kostyantyn Guzhva and Denilson Barbosa	105
<i>TextGraphs 2024 Shared Task on Text-Graph Representations for Knowledge Graph Question Answering</i> Andrey Sakhovskiy, Mikhail Salnikov, Irina Nikishina, Aida Usmanova, Angelie Kraft, Cedric Möller, Debayan Banerjee, Junbo Huang, Longquan Jiang, Rana Abdullah, Xi Yan, Dmitry Ustalov, Elena Tutubalina, Ricardo Usbeck and Alexander Panchenko	116
<i>nlp_enjoyers at TextGraphs-17 Shared Task: Text-Graph Representations for Knowledge Graph Question Answering using all-MPNet</i> Nikita Kurdiukov, Viktoriia Zinkovich, Sergey Karpukhin and Pavel Tikhomirov	126
<i>HW-TSC at TextGraphs-17 Shared Task: Enhancing Inference Capabilities of LLMs with Knowledge Graphs</i> Wei Tang, Xiaosong Qiao, Xiaofeng Zhao, Min Zhang, Chang Su, Yuang Li, Yinglu Li, Yilun Liu, Feiyu Yao, Shimin Tao, Hao Yang and He Xianghui	131
<i>TIGFORMER at TextGraphs-17 Shared Task: A Late Interaction Method for text and Graph Representations in KBQA Classification Task</i> Mayank Rakesh, Parikshit Saikia and Saket Kumar Shrivastava	137
<i>NLPeople at TextGraphs-17 Shared Task: Chain of Thought Questioning to Elicit Decompositional Reasoning</i> Movina Moses, Vishnudev Kuruvanthodi, Mohab Elkaref, Shinnosuke Tanaka, James Barry, Geeth De Mel and Campbell D Watson	142

Skoltech at TextGraphs-17 Shared Task: Finding GPT-4 Prompting Strategies for Multiple Choice Questions
Maria Lysyuk and Pavel Braslavski 149

JellyBell at TextGraphs-17 Shared Task: Fusing Large Language Models with External Knowledge for Enhanced Question Answering
Julia Belikova, Evgeniy Beliakin and Vasily Konovalov 154

Program

Thursday, August 15, 2024

- 09:00 - 09:10 *Opening Remarks*
- 09:10 - 09:30 *Learning Human Action Representations from Temporal Context in Lifestyle Vlogs*
- 09:30 - 09:50 *ConGraT: Self-Supervised Contrastive Pretraining for Joint Graph and Text Embeddings*
- 09:50 - 10:10 *A Pipeline Approach for Parsing Documents into Uniform Meaning Representation Graphs*
- 10:10 - 10:30 *Financial Product Ontology Population with Large Language Models*
- 10:30 - 11:00 *Break 1*
- 11:00 - 11:20 *Prompt Me One More Time: A Two-Step Knowledge Extraction Pipeline with Ontology-Based Verification*
- 11:20 - 11:40 *Towards Understanding Attention-based Reasoning through Graph Structures in Medical Codes Classification*
- 11:40 - 12:00 *Leveraging Graph Structures to Detect Hallucinations in Large Language Models*
- 12:00 - 12:20 *Semantic Graphs for Syntactic Simplification: A Revisit from the Age of LLM*
- 12:30 - 14:00 *Lunch Break*
- 14:00 - 15:00 *Invited Talk: A Legacy of Graphs: Celebrating Dragomir Radev's Contributions to Graph-Based Natural Language Processing*
- 15:00 - 15:30 *TextGraphs 2024 Shared Task on Text-Graph Representations for Knowledge Graph Question Answering.*
- 15:30 - 16:00 *Break 2*
- 16:00 - 17:30 *Poster Session*