

KnowLLM 2024

**The 1st Workshop on Towards Knowledgeable Language  
Models**

**Proceedings of the Workshop**

August 16, 2024

The KnowLLM organizers gratefully acknowledge the support from the following sponsors.

**Gold**



©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](mailto:acl@aclweb.org)

ISBN 979-8-89176-154-4

## Introduction

Welcome to KnowLLM 2024, the inaugural workshop on knowledgeable language models. Co-located with ACL 2024, this workshop is scheduled for August 16, 2024 to be held in Bangkok, Thailand.

Knowledge has been an important prerequisite for a variety of NLP applications, and is typically sourced from either structured knowledge sources such as knowledge bases and dictionaries or unstructured knowledge sources such as Wikipedia documents. More recently, researchers have discovered that language models already possess a significant amount of knowledge through pretraining: LLMs can be used to generate commonsense knowledge and factual knowledge context for question answering. While the results are encouraging, there are still lingering questions: Where does this knowledge come from? How much do language models know? Is this knowledge reliable? If some knowledge is wrong, can we fix it?

In response to these questions, the KnowLLM workshop examines the lifecycle of knowledge within language models: (1) the emergence of knowledge through language model pre-training; (2) injection of external knowledge; (3) the updating and modification of knowledge; (4) probing and generation of knowledge. Currently, researchers that focus on different stages in this lifecycle are scattered across different sub-communities within NLP: probing knowledge and editing knowledge is often associated with the interpretability track while injecting knowledge is often application-specific and is discussed within the dialog, QA, IE, or summarization tracks. The workshop seeks to bring these researchers together and facilitate collaboration to create a more holistic view of the problem.

The KnowLLM workshop is also closely related to some of the core challenges involving LM research: reducing hallucination, improving interpretability, and making models extensible. Although such challenges are still open, it is clear that knowledge plays a key role: (1) attribution to sources or providing the relevant knowledge during generation can mitigate hallucination; (2) being able to locate and trace the knowledge provides insight into the LM’s inner workings; (3) being able to efficiently adapt to domain knowledge or integrate updated facts improves extensibility.

This year, there were a total of 78 archival and non-archival submissions to the KnowLLM workshop, of which a total of 48 were accepted. Among these works, 16 have been included in our proceedings and 19 are included in ACL Findings.

In addition to oral and poster sessions where accepted works will be presented, the Workshop also will also host talks and a panel discussion with six invited speakers: Isabelle Augenstein, Peter Clark, Tatsunori Hashimoto, Ed Hovy, Hannah Rashkin, and Luke Zettlemoyer.

Finally, we would like to express our gratitude to all the authors, committee members, invited speakers, and participants for helping make this workshop possible. We would also like to gratefully acknowledge our sponsor, Amazon, for their support.

# Organizing Committee

## Program Chairs

Sha Li, University of Illinois, Urbana-Champaign

Manling Li, Northwestern University

Michael J.Q. Zhang, University of Texas at Austin

Mor Geva, Tel Aviv University and Google Research

Peter Hase, University of North Carolina at Chapel Hill

Eunsol Choi, University of Texas at Austin

Heng Ji, University of Illinois, Urbana-Champaign

## Program Committee

### Reviewers

Yupeng Cao  
Hou Pong Chan  
Hung-Ting Chen  
Shiqi Chen  
Roi Cohen  
Farima Fatahi Bayat  
Shangbin Feng  
Yicheng Fu  
Benjamin Han  
Li Hao  
Jing Huang  
Yizhu Jiao  
Xiaomeng Jin  
Nora Kassner  
Jeonghwan Kim  
Yichuan Li  
Yanming Liu  
Yasumasa Onoe  
Cheng Qian  
Jiaming Shen  
Zhengliang Shi  
Dachun Sun  
Ziqi Wang  
Xiaozhi Wang  
Zhenhailong Wang  
Qineng Wang  
Qingyun Wang  
Zihao Wei  
Jinfeng Xiao  
Ruo Chen Xu  
Fangyuan Xu  
Yan Xu  
Sohee Yang  
Gal Yona  
Ori Yoran  
En Yu  
Zixuan Zhang  
Pingyue Zhang  
Yuji Zhang  
Ming Zhong  
Zexuan Zhong  
Sizhe Zhou

### Invited Speakers

Isabelle Augenstein, University of Copenhagen

Peter Clark, Allen Institute for AI  
Luke Zettlemoyer, University of Washington and Meta  
Ed Hovy, Carnegie Mellon University  
Tatsunori Hashimoto, Stanford  
Hannah Rashkin, Google Deepmind

## Table of Contents

<i>PhonologyBench: Evaluating Phonological Skills of Large Language Models</i> Ashima Suvarna, Harshita Khandelwal and Nanyun Peng .....	1
<i>Is Your Large Language Model Knowledgeable or a Choices-Only Cheater?</i> Nishant Balepur and Rachel Rudinger .....	15
<i>Reassess Summary Factual Inconsistency Detection with Large Language Model</i> Jiuding Yang, Hui Liu, Weidong Guo, Zhuwei Rao, Yu Xu and Di Niu .....	27
<i>Beyond Text: Unveiling Multimodal Proficiency of Large Language Models with MultiAPI Benchmark</i> Xiao Liu, Jianfeng Lin and Jiawei Zhang .....	32
<i>Retrieval-Augmented Knowledge Integration into Language Models: A Survey</i> Yuxuan Chen, Daniel Röder, Justus-Jonas Erker, Leonhard Hennig, Philippe Thomas, Sebastian Möller and Roland Roller .....	45
<i>ClinicalRAG: Enhancing Clinical Decision Support through Heterogeneous Knowledge Retrieval</i> Yuxing Lu, Xukai Zhao and Jinzhuo Wang .....	64
<i>Modeling Uncertainty and Using Post-fusion as Fallback Improves Retrieval Augmented Generation with LLMs</i> Ye Liu, Rui Meng, Meghana Moorthy Bhat, Shafiq Joty, Caiming Xiong, Yingbo Zhou and Semih Yavuz .....	69
<i>AcKnowledge: Acquired Knowledge Representation by Small Language Model Without Pre-training</i> Sourav Das, Sanjay Chatterji and Imon Mukherjee .....	83
<i>Knowledge Acquisition through Continued Pretraining is Difficult: A Case Study on r/AskHistorians</i> Jan Vincent Hoffbauer, Sylwester Sawicki, Marc Lenard Ulrich, Tolga Buz, Konstantin Dobler, Moritz Schneider and Gerard De Melo .....	96
<i>Beyond Probabilities: Unveiling the Misalignment in Evaluating Large Language Models</i> Chenyang Lyu, Minghao Wu and Alham Fikri Aji .....	109
<i>PromptRE: Weakly-Supervised Document-Level Relation Extraction via Prompting-Based Data Programming</i> Chufan Gao, Xulin Fan, Jimeng Sun and Xuan Wang .....	132
<i>Patent Response System Optimised for Faithfulness: Procedural Knowledge Embodiment with Knowledge Graph and Retrieval Augmented Generation</i> Jung-Mei Chu, Hao-Cheng Lo, Jieh Hsiang and Chun-Chieh Cho .....	146
<i>Safe-Embed: Unveiling the Safety-Critical Knowledge of Sentence Encoders</i> Jinseok Kim, Jaewon Jung, Sangyeop Kim, Sohhyung Park and Sungzoon Cho .....	156
<i>Measuring the Inconsistency of Large Language Models in Preferential Ranking</i> Xiutian Zhao, Ke Wang and Wei Peng .....	171
<i>Retrieval-augmented generation in multilingual settings</i> Nadezhda Chirkova, David Rau, Hervé Déjean, Thibault Formal, Stéphane Clinchant and Vassilina Nikoulina .....	177
<i>Retrieve, Generate, Evaluate: A Case Study for Medical Paraphrases Generation with Small Language Models</i> Ioana Buhnila, Aman Sinha and Mathieu Constant .....	189



# Program

## Friday, August 16, 2024

09:00 - 09:10	<i>Welcome Session</i>
09:10 - 09:45	<i>Invited Talk 1</i>
09:45 - 10:20	<i>Invited Talk 2</i>
10:20 - 10:50	<i>Coffee Break</i>
10:50 - 11:25	<i>Invited Talk 3</i>
11:25 - 12:30	<i>Oral Presentations</i>
12:30 - 13:30	<i>Lunch Break</i>
13:30 - 14:05	<i>Invited Talk 4</i>
14:05 - 14:40	<i>Invited Talk 5</i>
14:40 - 15:15	<i>Invited Talk 6</i>
15:15 - 16:00	<i>Panel Discussion</i>
16:00 - 17:30	<i>Poster Session</i>