

AKBC 2026

**The 11th Workshop on Automated Knowledge Base
Construction**

Proceedings of the Workshop

October 28, 2026

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

Diamond Level



Bloomberg
Engineering

©2026 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 979-8-89176-464-4

Preface

These proceedings present the contributions to the 11th Workshop on Automated Knowledge Base Construction (AKBC 2026), co-located with EMNLP 2026 and held on October 28, 2026 in Budapest, Hungary. AKBC brings together researchers and practitioners working on the construction, integration, and use of structured knowledge in the era of large language models (LLMs). While LLMs have revolutionized natural language processing, they remain prone to hallucinations, reasoning “mode-collapse” in open-ended generation, and a lack of factual provenance. Knowledge Bases (KBs) address a key missing piece of the generative era: They serve as ground truth for fact verification, the semantic backbone for constrained decoding in generation, and a resource behind Retrieval-Augmented Generation (RAG). AKBC sits at the intersection of natural language processing, knowledge representation, databases, and machine learning, with a particular focus on how symbolic structure can ground, constrain, and enhance generative models, both at training time and at inference time.

AKBC 2026 is centered around a distinguished lineup of invited speakers who provide the intellectual core of the workshop: Denny Vrandečić (Wikimedia Foundation), founder of Wikidata and co-creator of Semantic MediaWiki, who has shaped the landscape of collaborative knowledge bases and the Semantic Web; Mausam (IIT Delhi and University of Washington), an AAAI Fellow whose pioneering work spans information extraction, neuro-symbolic AI, and LLM agents for knowledge and reasoning; Heng Ji (University of Illinois Urbana-Champaign), a Grainger College of Engineering professor and ACL Fellow known for her work on multimedia multilingual information extraction, knowledge-enhanced LLMs, and her leadership of the NIST TAC Knowledge Base Population task; Alon Halevy (Google), an ACM Fellow recognized for his contributions to data integration and knowledge representation, whose early work on web data management and schema matching laid foundations for modern knowledge extraction; Dan Roth (University of Pennsylvania/Oracle), an ACL Fellow whose research integrates machine learning and reasoning for natural language understanding; Amir Globerson (Tel Aviv University/Google), whose research spans structured prediction, deep learning, and language and visual understanding; and Parisa Kordjamshidi (Michigan State University), whose pioneering research on spatial language grounding and multimodal LLMs bridges language understanding with visual and spatial reasoning. The workshop is designed to facilitate exchange between these senior researchers and the next generation of researchers presenting poster sessions, creating an intimate and collegial setting for dialogue across career stages.

The workshop follows a successful series of previous editions: as an independent conference in 2022, 2021, 2020, 2019; as a workshop at NeurIPS in 2017, NAACL in 2016, and NeurIPS in 2014; at CIKM in 2013 and NAACL in 2012; and as a stand-alone event in Grenoble, France in 2010.

We received 58 research and vision submissions and accepted 27 papers (23 direct submissions and 4 ARR commitments), an acceptance rate of 46.6%. The program combines invited talks, featured paper presentations, a poster session, a shared task session, and community discussion.

The shared task asked participants to predict the complete set of correct object strings for given subject – relation pairs, going beyond standard factual question answering to construct complete and precise knowledge base entries. Seventeen teams describe their systems in these proceedings. The winning teams were: “Coverage Is Not Knowledge: Why Closed-Book Knowledge Base Construction Looks Selection Bound” by Maksim Silchenko (winner); “MAVE: Multi-Angle Voting for Emission in Closed-Book Knowledge Base Construction” by Jonas Kunze and Andreas Both (second place); and “The Instruction Tax on Factual Recall: Base Models Can Beat Their Instruction-Tuned Counterparts” by Christophe Khalil and Aya Mourad (third place). We congratulate all three teams and thank everyone who took part!

We also gratefully acknowledge the support of our sponsors, SLB (formerly Schlumberger) and Bloomberg, whose generous contribution made this workshop possible.

Finally, we thank all authors for their contributions, the members of the program committee for their careful reviews and constructive feedback, and all participants for advancing research on knowledge base construction and knowledge-grounded generation.

The AKBC Organizing Committee

Jan-Christoph Kalo, University of Amsterdam, Netherlands

Ndapa Nakashole, University of California, San Diego, USA

Simon Razniewski, ScaDS.AI Dresden/Leipzig & TU Dresden, Germany

Fabian M. Suchanek, Télécom Paris, Institut Polytechnique de Paris, France

Andreas Vlachos, University of Cambridge, United Kingdom

Andrew McCallum, University of Massachusetts Amherst, USA

Organizing Committee

Workshop Chairs

Jan-Christoph Kalo, University of Amsterdam, Netherlands
Ndapa Nakashole, University of California, San Diego, USA
Simon Razniewski, ScaDS.AI Dresden/Leipzig & TU Dresden, Germany
Fabian M. Suchanek, Télécom Paris, Institut Polytechnique de Paris, France
Andreas Vlachos, University of Cambridge, United Kingdom
Andrew McCallum, University of Massachusetts Amherst, USA

Shared Task Organizers

Jan-Christoph Kalo, University of Amsterdam
Tuan-Phong Nguyen, Vietnam National University
Simon Razniewski, ScaDS.AI & TU Dresden
Bohui Zhang, King's College London

Program Committee

Program Committee

Sameer Singh, University of California, Irvine
Gerhard Weikum, Max Planck Institute for Informatics
Sunita Sarawagi, IIT Bombay
Annalisa Gentile, IBM
Nitisha Jain, Microsoft Research Cambridge
Sören Auer, University of Hannover
Danqi Chen, Princeton University
Doug Downey, Allen Institute for AI / Northwestern University
Zifeng Ding, Cambridge University
Moy Yuan, Amazon
Paolo Papotti, EURECOM
Gerard de Melo, Hasso Plattner Institute
Steffen Staab, University of Stuttgart
Rainer Gemulla, University of Mannheim
Luis Galárraga, INRIA Rennes
Lihu Chen, Imperial College London
Advitya Gemawat, Microsoft Research
Estevam Hruschka, Megagon Labs
Jingbo Shang, University of California, San Diego
Roi Cohen, HPI
Hiba Arnaout, Darmstadt University
Shrestha Ghosh, Tübingen University
Paul Groth, University of Amsterdam
Filip Ilievski, Vrije Universiteit Amsterdam
Russa Biswas, Aalborg University
Derry Wijaya, Monash University Indonesia
Pierre-Henri Paris, Paris-Saclay University
Daniel Daza, Vrije Universiteit Amsterdam
Jieying Chen, Vrije Universiteit Amsterdam
Selene Báez Santamaría, University of Zurich
Mohamed Gad-Elrab, Bosch Research
Benno Kruit, Amsterdam UMC
Andrea Schimmenti, Università di Bologna
Roman Klinger, University of Bamberg
Jay Pujara, USC Information Sciences Institute
Zachary Ives, University of Pennsylvania
Samy Haffoudhi, Institut Polytechnique de Paris
François Crespín, Institut Polytechnique de Paris
Christian Bizer, University of Mannheim
Aparna Varde, University of Southern Denmark
Wassila Ouerdane, CentraleSupélec, Paris-Saclay University
Raphaël Troncy, EURECOM
Jonathan Berant, Tel-Aviv University/Google DeepMind
David Campos, University of Amsterdam
Robin Jia, University of Southern California
Sai Kolapudi, Amazon Web Services

Table of Contents

<i>Epistemic Twins: Enabling a Symbolic Science of Language Model Knowledge</i> Simon Razniewski, Shrestha Ghosh, Luca Giordano, Yujia Hu, Josua Kowalzik and Tuan-Phong Nguyen	1
<i>Auditing a KB Elicitation of Frontier LLM Knowledge: A Multi-dimensional Analysis of GPTKB v1.5</i> Shrestha Ghosh, Luca Giordano, Yujia Hu, Tuan-Phong Nguyen and Simon Razniewski	8
<i>Edge Knowledge Bases for Sustainable and Explainable AI</i> Aparna S. Varde, Emil Njor, Francesco Daghero and Mahyar Turchi Moghaddam	19
<i>A Temporal Knowledge Graph for Music Festival Lineup Forecasting</i> Julia Gastinger, Thilo Dieing, Christian Meilicke and Heiner Stuckenschmidt	26
<i>The Canonical Order Problem: When Large Language Models Are Unreliable Knowledge Bases for Multi-Valued Relations</i> Timo Pierre Schrader, Annemarie Friedrich, Simon Razniewski and Lukas Lange	45
<i>FALCON: A Model and Dataset Agnostic Framework for Synthetic Data Generation for NL2SQL Pairs</i> Darian Lee, Shannon Rumsey, Jack St. Clair, Xinyi Tang, Aditya Bansal and Yuanming Shi ..	56
<i>When Relevance Is Not Evidence: A Controlled Diagnostic of Claim-Conditioned Support Ranking</i> Ke Lu	88
<i>Fine-Grained Table Retrieval Through the Lens of Complex Queries</i> Wojciech Kosiuk, Xingyu Ji, Yeounoh Chung, Fatma Özcan and Madelon Hulsebos	94
<i>EventSocraticKG: Event-Centric Knowledge Graph Construction for Narrative Retrieval-Augmented Generation</i> William Rosengren	108
<i>HTTP for Meaning: Towards Linguistically Complete Knowledge Representations with Knowledge State Protocol</i> Ian Yuen	128
<i>When String Matching Stands In for a Knowledge Base: A Role-Aware Audit of WikiBigEdit</i> Edison Yang, Neel Marripalapu and Akshaj Agadi	139
<i>MUSART: A Scalable Multi-Domain Benchmark for Measuring Factual Recall in Large Language Models</i> Rajaa El Hamdani, Thomas Bonald and Fragkiskos Malliaros	151
<i>Scaling E-Commerce Attribute Extraction with Parallel Decoding</i> Nikhita Vedula, Dushyanta Dhyani, Bryan Wang and Shervin Malmasi	169
<i>Evidence-Carrying Claims: Fail-Closed Verification Gates for Language-Model-Assisted Research as Knowledge Base Construction</i> Robert Sneiderman	178
<i>Evidence-Guided Schema Normalization for Temporal Tabular Reasoning</i> Ashish Thanga, Vibhu Dixit, Abhilash Shankarampetta and Vivek Gupta	185
<i>GRACE: Graph-Grounded Reflective Agent Copilot Engine for Expert-in-the-Loop Knowledge Expansion</i> John Seon Keun Yi, Joshua R. Minot and Dokyun Lee	224

<i>Can We Do Interpretable NLI with Graphs Based on Atomic Propositions?</i>	
Younes Boufouss, Luc Pommeret, Thomas Gerald, Patrick Paroubek and Sophie Rosset	234
<i>A Rhetorical Citation Graph Model for Characterizing Scientific Findings: Resource Construction and Application to Plant Health</i>	
Anne-Sophie Foussat, Robert Bossy, Vincent Guigue, Nicolas Sauvion and Claire Nédellec .	247
<i>Schema-Guided Extraction of Interaction Knowledge Bases from Narrative Text: A Folktale Case Study</i>	
Sergey Kovalchuk, Anastasia Ugleva and Oleg Kubryak	261
<i>Operationalization Is Part of the Claim: Operationalization-Aware Knowledge in the Era of Generated Answers</i>	
Jan-Christoph Kalo	284
<i>GRACE-Mem: Graph Retrieval with Agentic Curation of Evidence for Long-Term Conversational Memory</i>	
Heng Wang, Xuan Ni Chen, Pei-Yun Chang, Chih-Hsi Chen, Hsin-Yu Huang, Chienyi Lin and I-Ling Chung	293
<i>Explaining Textual Entailment with Lexical Entailments: Using LLMs to Supply Lexical Relations for Formal Proofs</i>	
Jorryt de Jong, Stefan Morača, Ettore Cesari and Lasha Abzianidze	316
<i>LMKG: a flexible method for ontology-aware Knowledge Graph Extraction from text</i>	
Gianmarco Pappacoda and Paolo Torroni	334
<i>Zero-Shot Relation Classification with Large Language Models: Binary Prompting and Relevance Elicitation</i>	
Sheila Alemany, Danielle Sullivan-Pao and Olga Simek	343
<i>Distill-RE: Few-Shot Relation Extraction via Teacher LLM Reasoning Distillation</i>	
Danielle Sullivan-Pao, Sheila Alemany and Olga Simek	359
<i>From Imitation to Reasoning: Sample-Efficient LLM Agents for Knowledge Graph Completion</i>	
Chenjun Zheng, Zhigang Wang, Wenjia Zhang, Zhe Wang and Xiaolin Yang	379
<i>Revisiting RAG Retrievers: An Information Theoretic Benchmark</i>	
Wenqing Zheng, Dmitri Kalaev, Noah Fatsi, Daniel Barcklow, Owen Reinert, Igor Melnyk, Senthil Kumar and C. Bayan Bruss	393
<i>Summary of the AKBC 2026 Shared Task: Language Models for Knowledge Base Construction (LM-KBC, 5th Edition)</i>	
Jan-Christoph Kalo, Simon Razniewski, Bohui Zhang and Tuan-Phong Nguyen	409
<i>Closed-Book Knowledge Base Construction with Calibrated Abstention and Numeric Self-Consistency</i>	
Tanvir Ahmed	417
<i>A Forced Answer Channel for Closed-Book Knowledge Base Construction at LM-KBC 2026</i>	
Yiqun Sun, Pengfei Wei and Lawrence B. Hsieh	427
<i>REAP: Relation-Aware Elicitation and Parsing for Closed-Book Knowledge Base Construction from LLMs</i>	
Thanh-Dan Bui, Thanh-Trung Do and Tuan-Phong Nguyen	434
<i>SetProbe: Description-First Relation-Wise Self-Consistency for Closed-Book Knowledge Base Construction</i>	
Jannic Alexander Cutura	441

<i>Coverage Is Not Knowledge: Why Closed-Book Knowledge Base Construction Looks Selection Bound</i> Maksim Silchenko	447
<i>Knowledge Base Construction from Language Models: Mechanisms and Calibrations for Multi-Object Relational Extraction under Parameter Budgets</i> Chengcheng Wei, Ze Chen, Jiewen Zheng, Jiarong He and Max Gao	456
<i>Pre-training Aware Directed Reasoning for Closed-Book Knowledge Base Construction</i> Mikołaj Tym	462
<i>Exemplars Primarily Shift Retrieval, Verification Instructions Can Shift Rejection: An Anatomy of Inference-Time Interventions for Closed-Book Knowledge Base Construction</i> Takashi Sasaki	473
<i>COVER-KBC: Evidence-Centric Adaptive Inference for Closed-Book Knowledge Base Construction</i> Quoc-Linh Vo	480
<i>MAVE: Multi-Angle Voting for Emission in Closed-Book Knowledge Base Construction</i> Jonas Kunze and Andreas Both	488
<i>MuseList: Confidence-Gated Multi-Surface Recall for Closed-Book Knowledge Base Construction</i> Adrian Zimmermann	496
<i>The Instruction Tax on Factual Recall: Base Models Can Beat Their Instruction-Tuned Counterparts</i> Christophe Khalil and Aya Mourad	518
<i>One Workflow Does Not Fit All Relations: A System Description for the AKBC 2026 LM-KBC Shared Task</i> Andreea-Elena Cochintele, Anca Dobrovăț and Sergiu Nisioi	527
<i>CoRe: Construct-versus-Recall Routing for Knowledge Base Construction Under a Parameter Budget</i> Fatih Naufal Habibillah and Filbert Aurelian Tjitarianata	535
<i>Ensembling Elicitation Strategies for Knowledge Base Construction</i> Vlad-Mădălin Ungureanu, Matei Neacșu and Ianis Belu	543
<i>DataMind at AKBC Shared Task 2026: Relation-Specific Multiple Prompting for Closed-Book Knowledge Base Completion</i> Tanaree Patcharatrakul and Peerapon Vateekul	550
<i>Relation-Specific Elicitation of a Base Model under a 32B Closed-Book Budget for LM-KBC 2026</i> Ruggero Marino Lazzaroni, Tim Booker, Pau Soler Valadés, Mathias Angermaier and Kirill Solovev	560