L2M2 2025

The First Workshop on Large Language Model Memorization (L2M2)

Proceedings of the Workshop

August 1, 2025

The L2M2 organizers gratefully acknowledge the support from the following sponsors.

In collaboration with

Google

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 979-8-89176-278-7

Introduction

The First Workshop on Large Language Model Memorization (L2M2), co-located with ACL 2025 in Vienna, brings together researchers studying the phenomenon of memorization in large language models from multiple perspectives.

Large language models (LLMs) are known to memorize their training data, and this phenomenon has inspired multiple distinct research directions. Some researchers focus on understanding LLM memorization, attempting to localize memorized knowledge or identify which examples are most likely to be memorized. Others aim to edit or remove information that an LLM has memorized. Still others study the downstream implications of LLM memorization, including legal concerns associated with memorizing copyrighted articles, privacy risks associated with LLMs leaking private information, and benchmarking concerns that LLMs are memorizing test data.

This workshop seeks to provide a central venue for researchers studying LLM memorization from these different angles, fostering collaboration and advancing our understanding of this important phenomenon.

Organizing Committee

Workshop Chair

Robin Jia, University of Southern California

Workshop Chair

Eric Wallace, OpenAI and University of California Berkeley

Workshop Chair

Yangsibo Huang, Google

Workshop Chair

Tiago Pimentel, ETH Zürich

Workshop Chair

Pratyush Maini, Carnegie Mellon University

Workshop Chair

Verna Dankers, University of Edinburgh

Workshop Chair

Johnny Wei, University of Southern California

Workshop Chair

Pietro Lesci, University of Cambridge

Program Committee

Program Chairs

Verna Dankers, University of Edinburgh
Yangsibo Huang, Google
Robin Jia, University of Southern California
Pietro Lesci, University of Cambridge
Pratyush Maini, Carnegie Mellon University
Tiago Pimentel, ETH Zürich
Johnny Wei, University of Southern California

Area Chairs

Verna Dankers, University of Edinburgh
Yangsibo Huang, Google
Robin Jia, University of Southern California
Pietro Lesci, University of Cambridge
Pratyush Maini, Carnegie Mellon University
Tiago Pimentel, ETH Zürich
Johnny Wei, University of Southern California

Reviewers

Aryaman Arora, Stanford University
Marco Bombieri, University of Verona
Peter Carragher, Carnegie Mellon University
Ting-Yun Chang, University of Southern California
Bowen Chen, University of Tokyo
Matthew Finlayson, University of Southern California
James Flemings, University of Southern California
Ameya Godbole, University of Southern California
Patrick Haller, Humboldt Universität Berlin
Skyler Hallinan, University of Southern California
Kanyao Han, Walmart Global Tech and University of Illinois at Urbana-Champaign
Yuzheng Hu, University of Illinois at Urbana-Champaign
Jing Huang, Stanford University
Shotaro Ishihara, Nikkei Inc.
Masaru Isonuma, National Institute of Informatics and Tohoku University
Matthew Jagielski, Google
Dongjun Jang, Seoul National University
Mohammad Aflah Khan
Cristina Lopes, University of California, Irvine
Lucie Charlotte Magister, University of Cambridge
Matthieu Meeus, Imperial College London
Marius Mosbach, McGill University and Mila - Quebec Artificial Intelligence Institute
Max Ploner, Humboldt Universität Berlin
Xiangyu Qi, Princeton University
Nishat Raihan, George Mason University

Leonardo Ranaldi, University of Edinburgh
Vikas Raunak, Google DeepMind
Suchir Salhan, University of Cambridge
Avi Schwarzschild, Carnegie Mellon University
Igor Shilov, Imperial College London
Anshuman Suri, Northeastern University
Anvith Thudi, University of Toronto
Martin Tutek, University of Zagreb
Juraj Vladika, Technische Universität München
Ryan Yixiang Wang, University of Southern California
Boyi Wei, Princeton University
Johnny Wei, University of Southern California
Fan Wu, University of Illinois
Chiyuan Zhang, Google
Bihe Zhao, CISPA Helmholtz Center for Information Security

Table of Contents