

📅 July 2-3, 2026

CONFERENCE WEBSITE: <https://gem-workshop.com/>

---

## | About the Conference

Event Type: Call for Papers

Conference: GEM at ACL 2026

Date: July 2nd or July 3rd, 2026

Location: San Diego, California, USA

Website: <https://gem-workshop.com/>

Contact: [gem-workshop-chairs@googlegroups.com](mailto:gem-workshop-chairs@googlegroups.com)

### Important Dates

 **JUL 02** CONFERENCE DATE  
**July 2-3, 2026**

## Overview

The fifth edition of the Natural Language Generation, Evaluation, and Metrics (GEM) Workshop will be at ACL 2026 in San Diego!

Evaluation of language models has grown to be a central theme in NLP research, while remaining far from solved. As LMs have become more powerful, errors have become tougher to spot and systems harder to distinguish. Evaluation practices are evolving rapidly—from living benchmarks like Chatbot Arena to LMs being used as evaluators themselves (e.g., LM as judge, autoraters). Further research is needed to understand the interplay between metrics, benchmarks, and human-in-the-loop evaluation, and their impact in real-world

settings

Topics of Interest

We welcome submissions related to, but not limited to, the following topics:

Automatic evaluation of generation systems, including the use of LMs as evaluators

Creating evaluation corpora, challenge sets, and living benchmarks

Critiques of benchmarking efforts, including contamination, memorization, and validity

Evaluation of cutting-edge topics in LM development, including long-context understanding, agentic capabilities, reasoning, and more

Evaluation as measurement beyond raw capability, including ideas such as robustness, reliability, and more

Multimodal evaluation across text, vision, and other modalities

Cost-aware and efficient evaluation methods applicable across languages and scenarios

Human evaluation and its role in the era of powerful LMs

Evaluation of sociotechnical systems employing large language models

Surveys and meta-assessments of evaluation methods, metrics, and benchmarks

Best practices for dataset and benchmark documentation

Industry applications of the above-mentioned topics, especially internal benchmarking or navigating the gap between academic metrics and real-world impact.

Special Tracks

Opinion and Statement Papers Track (New!)

We are introducing a special track for opinion and statement papers. These submissions will be presented in curated panel discussions, encouraging open dialogue on emerging topics in evaluation research.

We welcome bold, thought-provoking position papers that challenge conventional wisdom, propose new directions for the field, or offer critical perspectives on current evaluation practices. This track is an opportunity to spark discussion and debate—submissions need not present new empirical results but should offer well-argued viewpoints supported by scientific evidence (e.g. prior studies) that advance our collective thinking about evaluation.

ReproNLP

The ReproNLP Shared Task on Reproducibility of Evaluations in NLP has been run for six consecutive years (2021–2026). ReproNLP 2026 will be part of the GEM Workshop at ACL 2026 in

San Diego. It aims to (i) shed light on the extent to which past NLP evaluations have been reproducible, and (ii) draw conclusions regarding how NLP evaluations can be designed and reported in order to increase reproducibility.

Participants submit reports for their reproductions of human evaluations from previous NLP literature where they quantitatively assess the degree of reproducibility using methods described in Belz.

(2025). More details can be found in the first call for participation for RepronLP 2026 at <https://repronlp.github.io>.

### Workshop Format

We aim to organize the workshop in an inclusive, highly interactive, and discussion-driven format. Paper presentations will focus on themed poster sessions that allow presenters to interact with researchers from varied backgrounds and similar interests. The workshop will feature panels on emerging topics and multiple short keynotes by leading experts.

 GEM Comic-Con Edition!

In the spirit of San Diego's famous Comic-Con (July 23-26), this year's GEM will be a special Comic-Con edition! We encourage participants to embrace creativity! Whether that's through themed poster designs, comic-style slides, or dressing up as your favorite evaluation metric personified, we want this year's workshop to be memorable and fun!

### Submission Types

Submissions can take any of the following forms:

Archival Papers: Original and unpublished work, for

all the following tracks—Main, ReprONLP, and Opinion/Statement.

**Non-Archival Extended Abstracts:** Work already presented or under review at a peer-reviewed venue. This is an excellent opportunity to share recent or ongoing work with the GEM community without precluding future publication.

**Findings Papers:** We additionally welcome presentation of relevant papers accepted to Findings, and will share more information at a later date.

All accepted papers will be given up to an additional page to address reviewers comments.

#### Submission Guidelines

Papers to be reviewed should be submitted directly through OpenReview, selecting the appropriate track, and conform to ACL 2026 style guidelines

**Review requirement:** For each submitted paper, authors may be asked to provide 2 reviews (either one author doing 2 reviews, or two authors each doing one review)

**Length.**

Archival papers should be within 4–8 pages, and opinion/statement papers should be within 2–4 pages. We make no “Short” or “Long” paper distinctions; we advise authors to tailor their submission length proportional to their contribution.

Extended abstracts should be within 1-2 pages.

Opinion/Statement Papers: These should be titled with the "Position:" prefix.

Dual submission: Dual submission of archival papers is not allowed. Authors interested in presenting work submitted to a different venue should instead use the non-archival extended abstract track.

### Important Dates

March 19, 2026: Direct paper submission deadline

April 9, 2026: Pre-reviewed ARR commitment deadline

April 28, 2026: Notification of acceptance

May 14, 2026: Camera-ready paper due

June 4, 2026: Pre-recorded video due (hard deadline)

July 2-3, 2026: Workshop at ACL in San Diego

### Contact

For any questions, please check the workshop page or email the organisers: [gem-workshop-chairs@googlegroups.com](mailto:gem-workshop-chairs@googlegroups.com)

Research papers are invited in, but not limited to, the following areas:

Uncategorized

Artificial  
Intelligence &  
Machine Learning

---

© 2026 CallForPaper.org - All Rights Reserved

*Providing global research dissemination and event management services.*