

ACL 2026

**Beyond Alignment: Transdisciplinary Conversations on
Human-AI Futures**

Proceedings of the Workshop

July 4, 2026

©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 979-8-89176-419-4

From the Organizers

Welcome to the BATCH Workshop at ACL 2026.

From chatbots for text and image creation, to coding tools and the design of assistive “agents,” technologists leverage LLMs for systems that they strive to “align” with human values and needs. At the same time, research on the social impacts and possibilities of new AI implementations are increasingly central to scholarship in the humanities and social sciences. Yet, opportunities to connect these efforts are surprisingly rare—curtailing the chance for researchers across disciplines to share up-to-date knowledge and collaborate on new ideas. This workshop offers an opportunity for technologists, humanists, and social scientists to bridge differences and establish common ground, discuss work in progress, and map out an interdisciplinary agenda for future research. We are grateful to our fellow scholars who have agreed to share their ideas and join the conversation: Idris Abdulmumin, Emily M. Bender, Su Lin Blodgett, Michele Elam, David Leslie, and Michael McCarrin.

As organizers, we hope BATCH becomes more than a one day gathering. In fact, the workshop follows up an earlier meeting at Rutgers in August 2025, which brought researchers across disciplines to Rutgers under the auspices of the Critical AI@Rutgers Initiative, on the topic of Writing Education and Design-Oriented AI. One of the key lessons of that meeting was to problematize the frameworks and methods currently used in generative AI systems to pursue meaningful outcomes through language. Reinforcement learning methods, for example, often stumble on the complex objectives and fine-grained choices needed to support students who are learning to refine their writing practices.

The BATCH workshop continues this conversation. We hope it opens a space where people from different fields can think together with care, curiosity, and honesty about what AI is becoming and what we want it to become. Our goal is to leave with new questions, new collaborations, and the beginning of a shared agenda to help foster a broad community of practice—equally engaged with critical design perspectives and with technical challenges—to pursue new AI technologies collaboratively, inclusively, and in the public interest. We acknowledge the support of NSF Award 2427646 in making the workshop possible.

Organizing Committee

Organizing Committee

Malihe Alikhani, Northeastern University and Brookings
Camille Gagnier, Binghamton University and Critical AI
Lauren M. E. Goodlad, Rutgers–New Brunswick
Dan Roth, University of Pennsylvania and Oracle
Mark Sammons, Insight Global
Matthew Stone, Rutgers–New Brunswick

Program Committee

Program Committee

Malihe Alikhani, Northeastern University and Brookings
Camille Gagnier, Binghamton University and Critical AI
Lauren M. E. Goodlad, Rutgers–New Brunswick
Asteria Kaeberlein, Northeastern University
Cameron Domenico Kirk-Giannini, Rutgers–Newark
Saki Imai, Northeastern University
Dan Roth, University of Pennsylvania and Oracle
Mark Sammons, Insight Global
Matthew Stone, Rutgers–New Brunswick
Tom Williams, Colorado School of Mines

Table of Contents

“I Was a Young AI”: On Probing the Effectiveness of Intervening on Anthropomorphic AI System Outputs
Su Lin Blodgett, Myra Cheng and Alexandra Olteanu 1

Program

Saturday, July 4, 2026

09:00 - 10:30 *Values and Alignment*

Opening Remarks

Lauren M. E. Goodlad

The Umbrella Coup: Alignment, LLMs, and the Misdirection of Ethics Discourse in NLP

Clémentine Bleuze, Fanny Ducei, Karën Fort, Aurélie Névéol and Emily M. Bender

11:00 - 12:00 *Invited Session*

A Conversation on Humanities and AI

Michele Elam and Michael McCarrin

12:00 - 13:30 *Lunch*

13:30 - 15:30 *Ethics and Alignment*

“I Was a Young AI”: On Probing the Effectiveness of Intervening on Anthropomorphic AI System Outputs

Su Lin Blodgett, Myra Cheng and Alexandra Olteanu

Perspectives on Alignment from Low-Resource Languages

Idris Abdulmumin

Intelligence just ain't in the head: On misrecognition and mismeasurement in AGI discourse

David Leslie

16:00 - 18:00 *Discussion and White Paper Planning*