

NLP+CSS 2026

**The Seventh Workshop on Natural Language Processing and
Computational Social Science**

Proceedings of the Workshop

July 3, 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 979-8-89176-426-2

Introduction

Welcome to the Seventh Workshop on Natural Language Processing (NLP) and Computational Social Science (CSS)! This workshop continues a successful series, with many interdisciplinary contributions to make NLP techniques and insights standard practice in CSS research, as well as improve NLP through insights from the social sciences. We received 76 submissions and after a rigorous review process by our committee, we accepted 23 entries, 19 archival and 4 non-archival. We also hosted a shared task for the first time this year in partnership with the Opioid Industry Document Archive (OIDA), and we are delighted to include two shared task submissions in the proceedings. Our workshop program also includes keynote talks by three outstanding scholars: Philip Resnik, a professor in the Department of Linguistics and Institute for Advanced Computer Studies at the University of Maryland, Lucy Li, a professor in the Department of Computer Sciences at University of Wisconsin-Madison, and R. Stuart Geiger a professor at the University of California, San Diego with appointments in the Department of Communication and the Halıcıoğlu Data Science Institute.

We would like to thank the Program Committee members who reviewed the papers this year. They did a wonderful job providing high-quality reviews, and particularly helping with last minute emergency reviews. We would also like to thank the workshop participants for the opportunities to connect (or re-connect) and learn from each other.

Dallas Card, Anjalie Field, Katherine Keith, and Julia Mendelsohn (Co-Organizers)

Organizing Committee

Program Committee

Dallas Card, University of Michigan
Anjalie Field, Johns Hopkins University
Katherine Keith, Williams College / Cohere
Julia Mendelsohn, University of Maryland

Program Committee

Reviewers

Anurag Acharya
Daniel Acuna
Jisun An
Aparna Ananthasubramaniam
Maria Antoniak
Elliott Ash
Nicolas Audinet de Pieuchon
Blanca Calvo Figueras
Danqing Chen
Hong Chen
Hongyu Chen
Andres Cruz
Aida Davani
Felix Drinkall
Jonathan Dunn
Agnieszka Falenska
Mark Finlayson
Bijean Ghafouri
Andrew Halterman
Endre Hamerlik
Kevin S. Hawkins
Nathan Oken Hodas
Alexander Miserlis Hoyle
Weihang Huang
Abraham Israeli
Jonathan Ivey
Cheonkam Jeong
Connor Thomas Jerzak
Richard Johansson
Moa Johansson
Kristen Johnson
Kenneth Joseph
David Jurgens
Elma Kerz
Rasul Khanbayov
Philipp Koehn
Frauke Kreuter
Haewoon Kwak
Alexandria Leto
Carlo Lipizzi
Chang Liu
Li Lucy
Tessa Masis
Ben Miller
David Mimno
Joel Mire

Ashley Moran
Nazia Nafis
Jason K. Nam
Bill Noble
Brianna O'Boyle
Sebastian Padó
Thierry Poibeau
Prateek Puri
Sushrita Rakshit
Christopher Rashidian
Christoph Rauchegger
Manoel Horta Ribeiro
Anthony Rios
Martin Ruskov
Germans Savcisens
Djamé Seddah
Sadat Shahriar
Bangzhao Shu
Emily Silcock
Dan Simonson
Akshay Singh
Mingyao Song
Andreas Spitz
Ian Stewart
Martin Takáč
Samia Touileb
Takehito Utsuro
Gisela Vallejo
Vasudha Varadarajan
Rob Voigt
Charles Welch
Yinuo Xu
Jinghua Xu
Li Yue
Lechen Zhang
Yifan Zhang
Mian Zhong
Viktoria Zlomanova

Table of Contents

| | |
|---|-----|
| <i>Prompt Perturbations Reveal Human-Like Biases in Large Language Model Survey Responses</i> Jens Rupperecht, Georg Ahnert and Markus Strohmaier | 1 |
| <i>Borrowed Words, Borrowed Minds: Probing LLM Choice of English-Derived Loanwords in Japanese</i> Joseph James | 22 |
| <i>Does Local News Stay Local?: Online Content Shifts in Sinclair-Acquired Stations</i> Miriam Wanner, Sophia Hager and Anjalie Field | 37 |
| <i>Learning Moral Diversity: Modelling Individual Perspectives in Moral Classification of Texts</i> Yi Ren, Lewis Mitchell and Matthew Roughan | 83 |
| <i>Launch and Aftermath: Contrasting Social Media Responses to Chatbot Releases. The Cases of Meta’s Galactica and OpenAI’s ChatGPT</i> Maximilian Weber and Johannes B. Gruber | 95 |
| <i>When Do LLMs Need Human Experts? Evidence for Social Science from Jurisprudential Classification</i> Caroline Cheng, Edward Stiglitz, David Mimno and Matthew Wilkens | 103 |
| <i>An NLP Framework for Analyzing Corporate Strategic Behavior in the Opioid Industry Documents Archive</i> Duy Dang Phu and Thìn Đặng Văn | 113 |
| <i>Beyond Acoustics: Isolating Dialectal and Sociolinguistic Bias in Spanish ASR</i> Johnatan E. Bonilla | 123 |
| <i>Who Speaks for Whom? LLM-Generated Survey Data as a Proxy for Public Opinion</i> Radhakrishnan Venkatakrishnan, Travis Brodbeck and Michael D. Young | 133 |
| <i>Documenting Corporate Harm: A Semantic Action Trajectories Approach to the Opioid Industry Document Archive Shared Task</i> Ben Miller | 149 |
| <i>Toward Unsupervised Conceptual Metaphor Discovery: A Case Study in Online Immigration Discourse</i> Alexandria Leto and Maria Leonor Pacheco | 159 |
| <i>Simulating Social Attitudes with LLMs: Accuracy, Demographic Effects, and Refusal Behavior in the Sensitive Domain of Suicide Prevention</i> Cristina J. Perez, Michael P. Vasquez Jr, Philippe Giabbanelli and Patrick Y. Wu | 176 |
| <i>Gender Disparities in LLM-Based Intimate Partner Violence Detection</i> Tabia Tanzin Prama, Mikaela Irene Fudolig, Abigail M. Crocker, Christopher M. Danforth and Peter Dodds | 190 |
| <i>Datasets and Methods for Improving the Cultural Capabilities of NLP Systems: A Survey</i> Tania Chakraborty, Eylon Caplan, Zhaoqing Wu, Kevin Cushing, Bruce Qin, Shreya Havaldar and Dan Goldwasser | 198 |
| <i>Towards More Transparent Online Campaigning: Detecting Political Campaign Content in Election-related Social Media Posts</i> Abdullah Alabdullah, Conor Gaughan, Thomas Flavel, Shubhanjay Varma, Rachel Gibson, Marta Cantijoch, Alexandru Cernat and Riza Batista-Navarro | 249 |

| | |
|---|-----|
| <i>Mapping the Landscape of Unregulated eXplicit Contents on Reddit</i> Msvpj Sathvik, Manan Roy Choudhury, Rishita Agarwal, Sathwik Narkedimilli, Thao Ha, Liesel Sharabi and Vivek Gupta..... | 271 |
| <i>From Adoption to Adaptation: Tracing the Diffusion of New Emojis on Twitter</i> Yuhang Zhou, Xuan Lu and Wei Ai..... | 293 |
| <i>Social Construction of Urban Space: Using LLMs to Identify Neighborhood Boundaries From Craigslist Ads</i> Adam Visokay, Ruth Bagley, Chris Hess, Ian Kennedy, Kyle Crowder, Rob Voigt and Denis Peskoff..... | 307 |
| <i>The Hidden Language of Harm: Examining the Role of Emojis in Harmful Online Communication and Content Moderation</i> Yuhang Zhou, Yimin Xiao, Wei Ai and Ge Gao..... | 322 |