ComputEL 2025

Eight Workshop on the Use of Computational Methods in the Study of Endangered Languages

Proceedings of the Workshop

March 4-5, 2025

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

Gold

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 None

Introduction

These proceedings contain the papers presented at the 8th Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL-8), held on March 4–5, 2025 in Honolulu, Hawai‘i. The workshop is co-located with the 9th International Conference on Language Documentation & Conservation (ICLDC9) and offers hybrid attendance options, enabling participants to join either in-person or remotely.

As the name implies, this is the eighth workshop dedicated to the intersection of computational tools and endangered language research. The inaugural event took place at the Association for Computational Linguistics (ACL) main conference in Baltimore, Maryland in 2014. Subsequent workshops have been co-located with the International Conference on Language Documentation & Conservation at the University of Hawai‘i at Mānoa (2017, 2019, 2021, 2023) or ACL-related venues (2022 in Dublin, Ireland; 2024 in St. Julians, Malta). We are delighted to continue this tradition by returning to Honolulu, marking the fifth time the workshop has been held alongside ICLDC.

The primary aim of ComputEL-8 is to bring together computational researchers, documentary linguists, and community language practitioners. By uniting these diverse groups, the workshop fosters a collaborative environment for exchanging ideas, methods, and resources that support the documentation and revitalization of endangered languages. The organizers are gratified by the variety of contributions, reflecting the importance of collaborative efforts across different disciplines and communities.

This year, we received 45 submissions in the form of papers or extended abstracts. Following a thorough review process, 30 were accepted. In addition, 3 presentations formed our special session, titled “Building Tools Together,” which focused on strategies for joint development of language resources and technologies.

We extend our appreciation to all authors for their submissions and to the Program Committee for the thoughtful review of each proposal. We also thank the ICLDC9 organizers for their assistance in hosting this workshop. We hope that ComputEL-8 sparks discussions and partnerships that continue to enrich the field of endangered language research, ultimately contributing to more robust support for language communities worldwide.

Organizing Committee

General Chair

Jordan Lachler, University of Alberta

Deputy General Chair

Godfred Agyapong, University of Florida

Program Chairs

Antti Arppe, University of Alberta
Aditi Chaudhary, Google Research
Sarah Moeller, University of Florida
Shruti Rijhwani, Birla Institute of Technology and Science, Pilani
Daisy Rosenblum, University of British Columbia

Program Committee

Chairs

Jordan Lachler, University of Alberta
Godfred Agyapong, University of Florida
Antti Arppe, University of Alberta
Aditi Chaudhary, Google Research
Sarah Moeller, University of Florida
Shruti Rijhwani, Birla Institute of Technology and Science, Pilani
Daisy Rosenblum, University of British Columbia

Program Committee

Milind Agarwal, George Mason University
Felix Ameka, Leiden University Centre for Linguistics
Antonios Anastasopoulos, George Mason University
Gregory Anderson, Living Tongues Institute for Endangered Languages
Candy Angulo, SUNY at Buffalo
Helen Aristar-Dry, ICHL
Dorothee Beermann, Norwegian University of Science and Technology
Martin Benjamin, Kamusi Project International
Andrea Berez-Kroeker, University of Hawaii at Mānoa Department of Linguistics
Claire Bowern, Yale University
Katia Chirkova, CRLAO, CNRS
Rolando Coto-Solano, Dartmouth College
Christopher Cox, Carleton University
Anna Luisa Daigneault, Living Tongues Institute for Endangered Languages
Elizaveta Dorofeeva, Moscow State University
Suzanne Duncan, Te Hiku Media
Bill Dyer, University of Florida
Mengzhe Geng, National Research Council Canada
Luke Gessler, Indiana University
Jeff Good, University at Buffalo
Michael Wayne Goodman, University of Washington
Rachael Griffiths, EPHE
Njabulo Hadebe, Computational linguist
Christopher Hammerly, University of British Columbia
William N. Havard, LLL, Université d'Orléans and CNRS
Ryan Henke, University of Hawaii at Mānoa
Gary Holton, Univeristy of Hawai'i
David Huggins-Daines, Independent Researcher
Benjamin Hunt, George Mason University
Xin He Jiang, University of Victoria
Anagha Narasimha Joshi, University of Georgia, Athens, GA
Seth Katenkamp, Yale University
Anna Kazentseva, National Research Council of Canada
František Kratochvíl, Palacký University
Olga Kriukova, University of Saskatchewan
Roland Kuhn, unknown

Éric Le Ferrand, Boston College
Dylan Leddy, Boston College
Gianna Leoni, Te Reo Irirangi o Te Hiku o te Ika
Gina-Anne Levow, University of Washington
Patrick Littell, National Reseach Council Canada
Zoey Liu, University of Florida
Olga Lovick, University of Saskatchewan
Kavya Manohar, Digital University Kerala
Bradley McDonnell, University of Hawaii
Mel Mistica, Melbourne University
Timothy Montler, UNT
Emmanuel Ngue Um, University of Yaoundé 1
Ake Nicholas, University of Auckland
William O'Grady, University of Hawaii at Manoa
Maura O'Leary, Western Washington University
Shu Okabe, TUM
Aidan Pine, National Research Council Canada
Emily Prudhommeaux, Boston College
Robert Pugh, Indiana University
Manny Rayner, University of South Australia
Aleksandr Riaposov, Universität Hamburg
Enora Rice, University of Colorado
Elizabeth Salesky, Johns Hopkins University
Nay San, Stanford University
Emmanuel Schang, LLL, Univ. Orléans and CNRS
Yves Scherrer, University of Oslo
Katherine Schmirler, University of Lethbridge
Miikka Silfverberg, UBC
Mark Simmons, University of California San Diego
Gary Simons, SIL International
Sonal Sinha, Department of Linguistics, k.m. institute of hindi and linguistics, B.R.Ambedkar University
Conor Snoek, University of Lethbridge
Ngoc Tan Le, Industrial University of Ho Chi Minh, Université du Québec à Montréal
Nick Thieberger, The University of Melbourne
Maria Belen Ticona Oquendo, University of Buenos Aires
Jörg Tiedemann, University of Helsinki
Paul Trilsbeek, Max Planck Institute for Psycholinguistics
Trond Trosterud, UiT
Francis M. Tyers, Indiana University Bloomington
Daan van Esch, Google
Sahiinii Lemaina Veikho, University Of Bern, Switzerland
Nitin Venkateswaran, University of Florida
Sunny Walker, University of Hawaii at Hilo
Olivia Waring, University of Hawai'i Mānoa, Department of Linguistics
Jonathan Washington, Swarthmore College
Linda Wiechetek, UiT
Cheyenne Wing, University of Arizona
Winston Wu, University of Hawaii at Hilo
Fei Xia, University of Washington
Changbing Yang, University of British Columbia

Table of Contents