Second Workshop on South East Asian Language Processing 2025

Abu Dhabi, United Arab Emirates
20 January 2025

ISBN: 979-8-3313-1373-9

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2025) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2026)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

bAI-bAI: A Context-Aware Transliteration System for Baybayin Scripts
Jacob Simon D. Bernardo and Maria Regina Justina E. Estuar 1

NusaBERT: Teaching IndoBERT to be Multilingual and Multicultural
Wilson Wongso, David Samuel Setiawan, Steven Limcorn and Ananto Joyoadikusumo 10

Evaluating Sampling Strategies for Similarity-Based Short Answer Scoring: a Case Study in Thailand
Pachara Boonsarngsuk, Pacharapon Arpanantikul, Supakorn Hiranwipas, Wipu Watcharakajorn and Ekapol Chuangsuwanich 27

Thai Winograd Schemas: A Benchmark for Thai Commonsense Reasoning
phakphum artkaew 42

Anak Baik: A Low-Cost Approach to Curate Indonesian Ethical and Unethical Instructions
Sulthan Abiyyu Hakim, Rizal Setya Perdana and Tirana Noor Fatyanosa 52

Indonesian Speech Content De-Identification in Low Resource Transcripts
Rifqi Naufal Abdjul, Dessi Puji Lestari, Ayu Purwarianti, Candy Olivia Mawalim, Sakriani Sakti and Masashi Unoki 63

IndoMorph: a Morphology Engine for Indonesian
Ian Kamajaya and David Moeljadi 72

NusaDialogue: Dialogue Summarization and Generation for Underrepresented and Extremely Low-Resource Languages
Ayu Purwarianti, Dea Adhista, Agung Baptiso, Miftahul Mahfuzh, Yusrina Sabila, Aulia Adila, Samuel Cahyawijaya and Alham Fikri Aji 82