

# **Third Workshop on Simple and Efficient Natural Language Processing (SustainLP 2022)**

Abu Dhabi, United Arab Emirates  
7 December 2022

ISBN: 978-1-7138-9149-9

**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© (2022) by the Association for Computational Linguistics  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Efficient Two-Stage Progressive Quantization of BERT</i> Charles Le, Arash Ardakani, Amir Ardakani, Hang Zhang, Yuyan Chen, James J. Clark, Brett H. Meyer and Warren J. Gross .....	1
<i>KGRefiner: Knowledge Graph Refinement for Improving Accuracy of Translational Link Prediction Methods</i> Mohammad Javad Saeedizade, Najmeh Torabian and Behrouz Minaei-Bidgoli .....	10
<i>Algorithmic Diversity and Tiny Models: Comparing Binary Networks and the Fruit Fly Algorithm on Document Representation Tasks</i> Tanise Ceron, Nhut Truong and Aurelie Herbelot .....	17
<i>Look Ma, Only 400 Samples! Revisiting the Effectiveness of Automatic N-Gram Rule Generation for Spelling Normalization in Filipino</i> Lorenzo Jaime Yu Flores and Dragomir Radev .....	29
<i>Who Says Elephants Can't Run: Bringing Large Scale MoE Models into Cloud Scale Production</i> Young Jin Kim, Rawn Henry, Raffy Fahim and Hany Hassan .....	36
<i>Data-Efficient Auto-Regressive Document Retrieval for Fact Verification</i> James Thorne .....	44
<i>AfroLM: A Self-Active Learning-based Multilingual Pretrained Language Model for 23 African Languages</i> Bonaventure F. P. Dossou, Atnafu Lambebo Tonja, Oreen Yousuf, Salomey Osei, Abigail Oppong, Iyanuoluwa Shode, Oluwabusayo Olufunke Awoyomi and Chris Chinenye Emezue .....	52
<i>Towards Fair Dataset Distillation for Text Classification</i> Xudong Han, Aili Shen, Yitong Li, Lea Frermann, Timothy Baldwin and Trevor Cohn .....	65