

# **Workshop on Simple and Efficient Natural Language Processing (SustainLP 2021)**

Online  
10 November 2021

ISBN: 978-1-7138-9181-9

**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© (2021) 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>Low Resource Quadratic Forms for Knowledge Graph Embeddings</i> Zachary Zhou, Jeffery Kline, Devin Conathan and Glenn Fung .....	1
<i>Evaluating the carbon footprint of NLP methods: a survey and analysis of existing tools</i> Nesrine bannour, Sahar Ghannay, Aurélie Névéol and Anne-Laure Ligozat .....	11
<i>Limitations of Knowledge Distillation for Zero-shot Transfer Learning</i> Saleh Soltan, Haidar Khan and Wael Hamza .....	22
<i>Countering the Influence of Essay Length in Neural Essay Scoring</i> SungHo Jeon and Michael Strube .....	32
<i>Memory-efficient Transformers via Top-k Attention</i> Ankit Gupta, Guy Dar, Shaya Goodman, David Ciprut and Jonathan Berant .....	39
<i>BioCopy: A Plug-And-Play Span Copy Mechanism in Seq2Seq Models</i> Yi Liu, Guoan Zhang, Puning Yu, Jianlin Su and Shengfeng Pan .....	53
<i>Combining Lexical and Dense Retrieval for Computationally Efficient Multi-hop Question Answering</i> Georgios Sidiropoulos, Nikos Voskarides, Svitlana Vakulenko and Evangelos Kanoulas .....	58
<i>Learning to Rank in the Age of Muppets: Effectiveness–Efficiency Tradeoffs in Multi-Stage Ranking</i> Yue Zhang, ChengCheng Hu, Yuqi Liu, Hui Fang and Jimmy Lin .....	64
<i>Improving Synonym Recommendation Using Sentence Context</i> Maria Glenski, William I. Sealy, Kate Miller and Dustin Arendt .....	74
<i>Semantic Categorization of Social Knowledge for Commonsense Question Answering</i> Gengyu Wang, Xiaochen Hou, Diyi Yang, Kathleen McKeown and Jing Huang .....	79
<i>Speeding Up Transformer Training By Using Dataset Subsampling - An Exploratory Analysis</i> Lovre Torbarina, Velimir Mihelčić, Bruno Šarlija, Lukasz Roguski and Željko Kraljević .....	86
<i>Hyperparameter Power Impact in Transformer Language Model Training</i> Lucas Høyberg Puvlis de Chavannes, Mads Guldborg Kjeldgaard Kongsbak, Timmie Rantzau and Leon Derczynski .....	96
<i>Distiller: A Systematic Study of Model Distillation Methods in Natural Language Processing</i> Haoyu He, Xingjian Shi, Jonas Mueller, Sheng Zha, Mu Li and George Karypis .....	119
<i>Shrinking Bigfoot: Reducing wav2vec 2.0 footprint</i> Zilun Peng, Akshay Budhkar, Ilana Tuil, Jason Levy, Parinaz Sobhani, Raphael Cohen and Jumana Nassour .....	134
<i>On the Role of Corpus Ordering in Language Modeling</i> Ameeta Agrawal, Suresh Singh, Lauren Schneider and Michael Samuels .....	142
<i>Efficient Domain Adaptation of Language Models via Adaptive Tokenization</i> Vin Sachidananda, Jason Kessler and Yi-An Lai .....	155
<i>Unsupervised Contextualized Document Representation</i> Ankur Gupta and Vivek Gupta .....	166