

# **Fourth Workshop on Simple and Efficient Natural Language Processing (SustainLP 2023)**

Toronto, Canada  
13 July 2023

ISBN: 978-1-7138-8235-0

**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© (2023) 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>KwikBucks: Correlation Clustering with Cheap-Weak and Expensive-Strong Signals</i> Sandeep Silwal, Sara Ahmadian, Andrew Nystrom, Andrew Mccallum, Deepak Ramachandran and Mehran Kazemi .....	1
<i>Semantic-Oriented Unlabeled Priming for Large-Scale Language Models</i> Yanchen Liu, Timo Schick and Hinrich Schtze .....	32
<i>oBERTa: Improving Sparse Transfer Learning via improved initialization, distillation, and pruning regimes</i> Daniel Campos, Alexandre Marques, Mark Kurtz and Cheng Xiang Zhai .....	39
<i>Quick Dense Retrievers Consume KALE: Post Training KullbackLeibler Alignment of Embeddings for Asymmetrical dual encoders</i> Daniel Campos, Alessandro Magnani and Chengxiang Zhai .....	59
<i>Lessons on Parameter Sharing across Layers in Transformers</i> Sho Takase and Shun Kiyono .....	78
<i>To Asymmetry and Beyond: Structured Pruning of Sequence to Sequence Models for Improved Inference Efficiency</i> Daniel Campos and Chengxiang Zhai .....	91
<i>Small is the New Big: Pre-finetuned compact models are better for Asynchronous Active Learning</i> Dantong Liu, Kaushik Pavani and Sunny Dasgupta .....	110
<i>ADEPT: Adapter-based Efficient Prompt Tuning Approach for Language Models</i> Aditya Shah, Surendrabikram Thapa, Aneesh Jain and Lifu Huang .....	121
<i>NLU on Data Diets: Dynamic Data Subset Selection for NLP Classification Tasks</i> Jean-michel Attendu and Jean-philippe Corbeil .....	129
<i>On the Interactions of Structural Constraints and Data Resources for Structured Prediction</i> Zhisong Zhang, Emma Strubell and Eduard Hovy .....	147
<i>Can we Pretrain a SotA Legal Language Model on a Budget From Scratch?</i> Joel Niklaus and Daniele Giofre .....	158
<i>Is a Video worth <math>n \times n</math> Images? A Highly Efficient Approach to Transformer-based Video Question An- swering</i> Chenyang Lyu, Tianbo Ji, Yvette Graham and Jennifer Foster .....	183
<i>How to Unleash the Power of Large Language Models for Few-shot Relation Extraction?</i> Xin Xu, Yuqi Zhu, Xiaohan Wang and Ningyu Zhang .....	190
<i>Prompting language models improves performance in imbalanced setting</i> Jay Mohta .....	201
<i>KGQA Without Retraining</i> Nick Mckenna and Priyanka Sen .....	212
<i>MANER: Mask Augmented Named Entity Recognition for Extreme Low-Resource Languages</i> Shashank Sonkar, Zichao Wang and Richard Baraniuk .....	219

<i>Efficient and Interpretable Compressive Text Summarisation with Unsupervised Dual-Agent Reinforcement Learning</i>	
Peggy Tang, Junbin Gao, Lei Zhang and Zhiyong Wang .....	227
<i>Exploring the Effect of Frequency Resolution in FNet</i>	
Gregory Szumel, Ghazal Khalighinejad, Rickard Stureborg and Sam Wiseman .....	239
<i>Towards Adaptable and Interactive Image Captioning with Data Augmentation and Episodic Memory</i>	
Aliki Anagnostopoulou, Mareike Hartmann and Daniel Sonntag .....	245
<i>Corpus Complexity Matters in Pretraining Language Models</i>	
Ameeta Agrawal and Suresh Singh .....	257
<i>PersonaPKT: Building Personalized Dialogue Agents via Parameter-efficient Knowledge Transfer</i>	
Xu Han, Bin Guo, Yoon Jung, Benjamin Yao, Yu Zhang, Xiaohu Liu and Chenlei Guo .....	264
<i>Small Character Models Match Large Word Models for Autocomplete Under Memory Constraints</i>	
Ganesh Jawahar, Subhabrata Mukherjee, Debadeepta Dey, Muhammad Abdul-mageed, Laks Lakshmanan, V.s., Caio Mendes, Gustavo De Rosa and Shital Shah .....	274
<i>Query Encoder Distillation via Embedding Alignment is a Strong Baseline Method to Boost Dense Retriever Online Efficiency</i>	
Yuxuan Wang and Lyu Hong .....	290
<i>Minimalist Entity Disambiguation for Mid-Resource Languages</i>	
Benno Kruit .....	299