

# **Fourth Workshop on Insights from Negative Results in NLP (Insights 2023)**

Dubrovnik, Croatia  
5 May 2023

ISBN: 978-1-7138-7383-9

**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© (2023) by the Association for Computational Linguistics  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2023)

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>Missing Information, Unresponsive Authors, Experimental Flaws: The Impossibility of Assessing the Reproducibility of Previous Human Evaluations in NLP</i>	
Anya Belz, Craig Thomson and Ehud Reiter . . . . .	1
<i>ERATE: Efficient Retrieval Augmented Text Embeddings</i>	
Vatsal Raina, Nora Kassner, Kashyap Popat, Patrick Lewis, Nicola Cancedda and Louis Martin	11
<i>A Data-centric Framework for Improving Domain-specific Machine Reading Comprehension Datasets</i>	
Iva Bojic, Josef Halim, Verena Suharman, Sreeja Tar, Qi Chwen Ong, Duy Phung, Mathieu Ravaut, Shafiq Joty and Josip Car . . . . .	19
<i>Encoding Sentence Position in Context-Aware Neural Machine Translation with Concatenation</i>	
Lorenzo Lupo, Marco Dinarelli and Laurent Besacier . . . . .	33
<i>SocBERT: A Pretrained Model for Social Media Text</i>	
Yuting Guo and Abeed Sarker . . . . .	45
<i>Edit Aware Representation Learning via Levenshtein Prediction</i>	
Edison Marrese-taylor, Machel Reid and Alfredo Solano . . . . .	53
<i>What changes when you randomly choose BPE merge operations? Not much.</i>	
Jonne Saleva and Constantine Lignos . . . . .	59
<i>Hiding in Plain Sight: Insights into Abstractive Text Summarization</i>	
Vivek Srivastava, Savita Bhat and Niranjan Pedanekar . . . . .	67
<i>Annotating PubMed Abstracts with MeSH Headings using Graph Neural Network</i>	
Faizan Mustafa, Rafika Boutalbi and Anastasiia Iurshina . . . . .	75
<i>Do not Trust the Experts: How the Lack of Standard Complicates NLP for Historical Irish</i>	
Oksana Dereza, Theodorus Franssen and John P. Mccrae . . . . .	82
<i>Exploring the Reasons for Non-generalizability of KBQA systems</i>	
Sopan Khosla, Ritam Dutt, Vinayshekhar Bannihatti Kumar and Rashmi Gangadharaiah . . . . .	88
<i>An Empirical Study on Active Learning for Multi-label Text Classification</i>	
Mengqi Wang and Ming Liu . . . . .	94