

24th Workshop on Biomedical Language Processing (BioNLP 2025)

Vienna, Austria
1 August 2025

ISBN: 979-8-3313-2402-5

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. (2025)

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

<i>Understanding the Impact of Confidence in Retrieval Augmented Generation: A Case Study in the Medical Domain</i>	
Shintaro Ozaki, Yuta Kato, Siyuan Feng, Masayo Tomita, Kazuki Hayashi, Wataru Hashimoto, Ryoma Obara, Masafumi Oyamada, Katsuhiko Hayashi, Hidetaka Kamigaito and Taro Watanabe	1
<i>Effect of Multilingual and Domain-adapted Continual Pre-training on Few-shot Promptability</i>	
Ken Yano and Makoto Miwa	18
<i>MedSummRAG: Domain-Specific Retrieval for Medical Summarization</i>	
Guanting Luo and Yuki Arase	27
<i>Enhancing Stress Detection on Social Media Through Multi-Modal Fusion of Text and Synthesized Visuals</i>	
Efstathia Soufleri and Sophia Ananiadou	34
<i>Fine-tuning LLMs to Extract Epilepsy Seizure Frequency Data from Health Records</i>	
Ben Holgate, Joe Davies, Shichao Fang, Joel Winston, James Teo and Mark Richardson	44
<i>AdaBioBERT: Adaptive Token Sequence Learning for Biomedical Named Entity Recognition</i>	
Sumit Kumar and Tanmay Basu	56
<i>Transformer-Based Medical Statement Classification in Doctor-Patient Dialogues</i>	
Farnod Bahrololloomi, Johannes Luderschmidt and Biying Fu	63
<i>PreClinIE: An Annotated Corpus for Information Extraction in Preclinical Studies</i>	
Simona Doneva, Hanna Hubarava, Pia Härvelid, Wolfgang Zürrer, Julia Bugajska, Bernard Hild, David Brüscheiler, Gerold Schneider, Tilja Ellendorff and Benjamin Ineichen	74
<i>Benchmarking zero-shot biomedical relation triplet extraction across language model architectures</i>	
Frederik Gade, Ole Lund and Marie Lisandra Mendoza	88
<i>RadQA-DPO: A Radiology Question Answering System with Encoder-Decoder Models Enhanced by Direct Preference Optimization</i>	
Md Sultan Al Nahian and Ramakanth Kavuluru	101
<i>Gender-Neutral Large Language Models for Medical Applications: Reducing Bias in PubMed Abstracts</i>	
Elizabeth Schaefer and Kirk Roberts	114
<i>Error Detection in Medical Note through Multi Agent Debate</i>	
Abdine Maiga, Anoop Shah and Emine Yilmaz	124
<i>Accelerating Cross-Encoders in Biomedical Entity Linking</i>	
Javier Sanz-Cruzado and Jake Lever	136
<i>Advancing Biomedical Claim Verification by Using Large Language Models with Better Structured Prompting Strategies</i>	
Siting Liang and Daniel Sonntag	148
<i>A Retrieval-Based Approach to Medical Procedure Matching in Romanian</i>	
Andrei Niculae, Adrian Cosma and Emilian Radoi	167
<i>Improving Barrett's Oesophagus Surveillance Scheduling with Large Language Models: A Structured Extraction Approach</i>	
Xinyue Zhang, Agathe Zecevic, Sebastian Zeki and Angus Roberts	176

<i>Prompting Large Language Models for Italian Clinical Reports: A Benchmark Study</i>	
Livia Lilli, Carlotta Masciocchi, Antonio Marchetti, Giovanni Arcuri and Stefano Patarnello	190
<i>QoLAS: A Reddit Corpus of Health-Related Quality of Life Aspects of Mental Disorders</i>	
Lynn Greschner, Amelie Wühl and Roman Klinger	201
<i>LLMs as Medical Safety Judges: Evaluating Alignment with Human Annotation in Patient-Facing QA</i>	
Yella Diekmann, Chase Fensore, Rodrigo Carrillo-Larco, Eduard Castejon Rosales, Sakshi Shiromani, Rima Pai, Megha Shah and Joyce Ho	217
<i>Effective Multi-Task Learning for Biomedical Named Entity Recognition</i>	
João Ruano, Gonalo Correia, Leonor Barreiros and Afonso Mendes	225
<i>Can Large Language Models Classify and Generate Antimicrobial Resistance Genes?</i>	
Hyunwoo Yoo, Haebin Shin and Gail Rosen	240
<i>CaseReportCollective: A Large-Scale LLM-Extracted Dataset for Structured Medical Case Reports</i>	
Xiao Yu Cindy Zhang, Melissa Fong, Wyeth Wasserman and Jian Zhu	249
<i>Enhancing Antimicrobial Drug Resistance Classification by Integrating Sequence-Based and Text-Based Representations</i>	
Hyunwoo Yoo, Bahrad Sokhansanj and James Brown	263
<i>Questioning Our Questions: How Well Do Medical QA Benchmarks Evaluate Clinical Capabilities of Language Models?</i>	
Siun Kim and Hyung-Jin Yoon	274
<i>Beyond Citations: Integrating Finding-Based Relations for Improved Biomedical Article Representations</i>	
Yuan Liang, Massimo Poesio and Roonak Rezvani	297
<i>Converting Annotated Clinical Cases into Structured Case Report Forms</i>	
Pietro Ferrazzi, Alberto Lavelli and Bernardo Magnini	307
<i>MuCoS: Efficient Drug–Target Discovery via Multi-Context-Aware Sampling in Knowledge Graphs</i>	
Haji Gul, Abdul Naim and Ajaz Bhat	319
<i>Overcoming Data Scarcity in Named Entity Recognition: Synthetic Data Generation with Large Language Models</i>	
An Dao, Hiroki Teranishi, Yuji Matsumoto, Florian Boudin and Akiko Aizawa	328
<i>PetEVAL: A veterinary free text electronic health records benchmark</i>	
Sean Farrell, Alan Radford, Noura Al Moubayed and Peter-John Noble	341
<i>Virtual CRISPR: Can LLMs Predict CRISPR Screen Results?</i>	
Steven Song, Abdalla Abdrabou, Asmita Dabholkar, Kastan Day, Pavan Dharmaju, Jason Perera, Volodymyr Kindratenko and Aly Khan	354
<i>Overview of the BioLaySumm 2025 Shared Task on Lay Summarization of Biomedical Research Articles and Radiology Reports</i>	
Chenghao Xiao, Kun Zhao, Xiao Wang, Siwei Wu, Sixing Yan, Tomas Goldsack, Sophia Ananiadou, Noura Al Moubayed, Liang Zhan, William K. Cheung and Chenghua Lin	365
<i>Overview of the ClinQLink 2025 Shared Task on Medical Question-Answering</i>	
Brandon Colelough, Davis Bartels and Dina Demner-Fushman	378

<i>SMAFIRA Shared Task at the BioNLP'2025 Workshop: Assessing the Similarity of the Research Goal</i>	
Mariana Neves, Iva Sovadinova, Susanne Fieberg, Celine Heint, Diana Rubel, Gilbert Schönfelder and Bettina Bert	388
<i>Overview of the ArchEHR-QA 2025 Shared Task on Grounded Question Answering from Electronic Health Records</i>	
Sarvesh Soni, Soumya Gayen and Dina Demner-Fushman	396