ClinicalNLP 2024

The 6th Workshop on Clinical Natural Language Processing (ClinicalNLP)

Proceedings of the Workshop

June 21, 2024

The ClinicalNLP organizers gratefully acknowledge the support from the following sponsors.

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-109-4

Preface

This volume contains papers from the 6th Workshop on Clinical Natural Language Processing (Clinical NLP), held at NAACL 2024.

Much of the information recorded in a clinical encounter is located exclusively in provider narrative notes, which makes them indispensable for supplementing structured clinical data in order to better understand patient state and care provided. The goal of this workshop is to bring together researchers interested in improving NLP technology to enable clinical applications, focusing on information extraction and modeling of narrative provider notes from electronic health records, patient encounter transcripts, and other clinical narratives. This year, we received a total of 48 submissions to the main workshop, of which 8 were accepted as oral presentations, and 21 were accepted as poster presentations.

ClinicalNLP 2024 also hosted four shared tasks, challenging researchers around the world to develop new approaches to solve clinical NLP problems: medical error detection and correction, multilingual and multimodal medical answer generation, text-to-SQL modeling, and chemotherapy timelines extraction. In addition to the four task description papers from the four shared task organizers, we received a total of 34 participant submissions to the shared tasks, of which 4 were accepted as oral presentations, and 30 were accepted as poster presentations.

Organizing Committee

General Chairs

Tristan Naumann, Microsoft Research
Asma Ben Abacha, Microsoft
Steven Bethard, University of Arizona
Kirk Roberts, UTHealth Houston
Danielle Bitterman, Harvard Medical School

MEDIQA-CORR 2024 Shared Task Organizers

Asma Ben Abacha, Microsoft
Wen-wai Yim, Microsoft
Meliha Yetisgen, University of Washington
Fei Xia, University of Washington

MEDIQA-M3G 2024 Shared Task Organizers

Asma Ben Abacha, Microsoft
Wen-wai Yim, Microsoft
Meliha Yetisgen, University of Washington
Fei Xia, University of Washington
Martin Krallinger, Barcelona Supercomputing Center (BSC)

EHRSQL 2024 Shared Task Organizers

Edward Choi, KAIST
Gyubok Lee, KAIST
Sunjun Kweon, KAIST
Seongsu Bae, KAIST

ChemoTimelines 2024 Shared Task Organizers

Jiarui Yao, Boston Children's Hospital, Harvard Medical School
Guergana Savova, Boston Children's Hospital, Harvard Medical School
Harry Hochheiser, University of Pittsburgh
WonJin Yoon, Boston Children's Hospital, Harvard Medical School
Eli Goldner, Boston Children's Hospital, Harvard Medical School

Program Committee

Program Chairs

Asma Ben Abacha, Steven Bethard, Danielle Bitterman, Tristan Naumann, Kirk Roberts

ChemoTimelines Reviewers

Nesrine Bannour, Asma Ben Abacha, Shohreh Haddadan, Yukun Tan, Liwei Wang, Wen-wai Yim, Xingmeng Zhao

EHRSQL Reviewers

Asma Ben Abacha, Satya Kesav Gundabathula, Sourav Bhowmik Joy, Sangryul Kim, Oleg Somov, Jerrin John Thomas, Wen-wai Yim

MEDIQA-CORR Reviewers

Jean-Philippe Corbeil, Aryo Pradipta Gema, Satya Kesav Gundabathula, Hyeon Hwang, Suramya Jadhav, Gyubok Lee, Juan Pajaro, Swati Rajwal, Nadia Saeed, Augustin Toma, Airat Valiev, Zhaolong Wu, Jiarui Yao

MEDIQA-M3G Reviewers

Marie Bauer, Ricardo Omar Chávez García, Gyubok Lee, Nadia Saeed, Jerrin John Thomas, Augustin Toma, Parth Vashisht, Jiarui Yao

Main Workshop Reviewers

Ashwag Alasmari, Emily Alsentzer, Ibtihel Amara, Hajer Ayadi, Leonardo Campillos-Llanos, Shan Chen, Hong-Jie Dai, Dmitriy Dligach, Richard Dufour, Jocelyn Dunstan, Naome A Etori, Matúš Falis, Yadan Fan, Aryo Pradipta Gema, Zelalem Gero, John Michael Giorgi, Natalia Grabar, Mei-Hua Hall, Ming Huang, Raphael Iyamu, Qiao Jin, Yoshinobu Kano, Yejin Kim, Yanis Labrak, Egoitz Laparra, Alberto Lavelli, Ulf Leser, Qiuhao Lu, Yuxing Lu, Diwakar Mahajan, Sérgio Matos, George Michalopoulos, Timothy A Miller, Aurélie Névéol, Ankur Padia, Satyajeet Raje, Pavithra Rajendran, Giridhar Kaushik Ramachandran, Frank Rudzicz, Ashwyn Sharma, Sonish Sivarajkumar, Arvind Krishna Sridhar, Dhananjay Srivastava, Behrad Taghibeyglou, Khushboo Thaker, Augustin Toma, Jinge Wu, Susmitha Wunnava, Dongfang Xu, Wen-wai Yim, WonJin Yoon, Paul Youssef, Weipeng Zhou

Keynote Talk

Dual Edges of Innovation: Risks and benefits of LLMs in LMICs

David Restrepo
Universidad del Cauca

Abstract: Large language models (LLMs) have emerged as transformative forces within artificial intelligence, heralding new capabilities in numerous sectors, including healthcare. Yet, the dialogue about their risks and their potential to widen social disparities, particularly in low-resource settings, remains insufficiently explored. In this keynote, we will dissect the evolution and fundamental principles of natural language processing (NLP), with a focus on the advent of transformative transformer models and their implications for fairness and bias.

We will start by outlining basic NLP concepts, progressively delving into how transformer models have reshaped our understanding of human-language machine interactions. This discussion will serve as a foundation to address the significant, yet often subtle, challenges of fairness and bias that are inherent in these models. The pervasive integration of advanced NLP technologies in clinical applications carries risks of perpetuating, or even exacerbating, existing biases which could profoundly affect patient care and outcomes.

The discourse will then shift to explore the advantages and practical applications of LLMs, with a focus on use cases in the Latin American context. Through specific examples, we will illustrate how LLMs can be leveraged to bridge language barriers and improve healthcare delivery in low-resource settings. Additionally, we will examine case studies from clinical settings across Latin America, highlighting the critical need for vigilance and the implementation of corrective measures to ensure these powerful tools serve all communities equitably.

Bio: David Restrepo is an Electronics and Communications Engineer and Data Scientist from Colombia, currently serving as a researcher at MIT Critical Data. He has also conducted significant research at the Laboratory for Computational Physiology at MIT, USA, and the University of Cauca in Colombia.

David's research is primarily focused on the application of machine learning in healthcare. He is particularly dedicated to addressing health inequalities and biases by developing methods for bias detection and de-biasing in medical images, text, and electronic health records (EHR) data. Additionally, he is actively involved in open data initiatives and events that aim to build capacity in the field.

His technical expertise includes efficient multimodal deep learning techniques that integrate medical images, textual data, and tabular datasets. Beyond his research, David is committed to mentoring and plays a pivotal role in organizing global datathons. These events promote collaborative data science and foster a diverse and interdisciplinary ecosystem in healthcare settings.

Table of Contents

Exploring Robustness in Doctor-Patient Conversation Summarization: An Analysis of Out-of-Domain SOAP Notes
Yu-Wen Chen and Julia Hirschberg 1

Efficient Medical Question Answering with Knowledge-Augmented Question Generation
Julien Khlaut, Corentin Dancette, Elodie Ferreres, Benani D. Alaedine, Herent Herent and Pierre Manceron 10

Gemini Goes to Med School: Exploring the Capabilities of Multimodal Large Language Models on Medical Challenge Problems & Hallucinations
Ankit Pal and Malaikannan Sankarasubbu 21

Retrieval augmented text-to-SQL generation for epidemiological question answering using electronic health records
Angelo Ziletti and Leonardo DAmbrosi 47

ClinicalMamba: A Generative Clinical Language Model on Longitudinal Clinical Notes
Zhichao Yang, Avijit Mitra, Sunjae Kwon and Hong yu 54

Working Alliance Transformer for Psychotherapy Dialogue Classification
Baihan Lin, Guillermo Cecchi and Djallel Bouneffouf 64

Building A German Clinical Named Entity Recognition System without In-domain Training Data
Siting Liang and Daniel Sonntag 70

DAIC-WOZ: On the Validity of Using the Therapist's prompts in Automatic Depression Detection from Clinical Interviews
Sergio Burdisso, Ernesto A. Reyes-Ramírez, Esaú Villatoro-tello, Fernando Sánchez-Vega, Adrian Pastor Lopez Monroy and Petr Motlicek 82

Parameter-Efficient Fine-Tuning of LLaMA for the Clinical Domain
Aryo Pradipta Gema, Pasquale Minervini, Luke Daines, Tom Hope and Beatrice Alex 91

A Multilevel Analysis of PubMed-only BERT-based Biomedical Models
Vicente Ivan Sanchez Carmona, Shanshan Jiang and Bin Dong 105

A Privacy-Preserving Corpus for Occupational Health in Spanish: Evaluation for NER and Classification Tasks
Claudio Aracena, Luis Miranda, Thomas Vakili, Fabián Villena, Tamara Quiroga, Fredy Núñez-Torres, Victor Rocco and Jocelyn Dunstan 111

DERA: Enhancing Large Language Model Completions with Dialog-Enabled Resolving Agents
Varun Nair, Elliot Schumacher, Geoffrey Jay Tso and Anitha Kannan 122

LlamaMTS: Optimizing Metastasis Detection with Llama Instruction Tuning and BERT-Based Ensemble in Italian Clinical Reports
Livia Lilli, Stefano Patarnello, Carlotta Masciocchi, Valeria Masiello, Fabio Marazzi, Tagliaferri Luca and Nikola Dino Capocchiano 162

Using Structured Health Information for Controlled Generation of Clinical Cases in French
Hugo Boulanger, Nicolas Hiebel, Olivier Ferret, Karën Fort and Aurélie Névéol 172

Large Language Models Provide Human-Level Medical Text Snippet Labeling
Ibtihel Amara, Haiyang Yu, Fan Zhang, Yuchen Liu, Benny LI, Chang Liu, Rupesh Kartha and Akshay Goel 185

Conversational Topic Recommendation in Counseling and Psychotherapy with Decision Transformer and Large Language Models
Aylin Ece Gunal, Baihan Lin and Djallel Bouneffouf 196

Leveraging Wikidata for Biomedical Entity Linking in a Low-Resource Setting: A Case Study for German
Faizan E Mustafa, Corina Dima, Juan G. Diaz Ochoa and Steffen Staab 202

Revisiting Clinical Outcome Prediction for MIMIC-IV
Tom Röhr, Alexei Figueroa, Jens-Michalis Papaioannou, Conor Fallon, Keno Bressem, Wolfgang Nejdl and Alexander Löser 208

Can LLMs Correct Physicians, Yet? Investigating Effective Interaction Methods in the Medical Domain
Burcu Sayin, Pasquale Minervini, Jacopo Staiano and Andrea Passerini 218

Leveraging pre-trained large language models for aphasia detection in English and Chinese speakers
Yan Cong, Jiyeon Lee and Arianna N. LaCroix 238

Fusion of Domain-Adapted Vision and Language Models for Medical Visual Question Answering
Cuong Ha, Shima Asaadi, Sanjeev Kumar Karn, Oladimeji Farri, Tobias Heimann and Thomas Runkler 246

LLM-Based Section Identifiers Excel on Open Source but Stumble in Real World Applications
Saranya Krishnamoorthy, Ayush Singh and Shabnam Tafreshi 258

Adapting Abstract Meaning Representation Parsing to the Clinical Narrative – the SPRING THYME parser
Jon Cai, Kristin Wright-Bettner, Martha Palmer, Guergana K Savova and James H. Martin . . 271

SERPENT-VLM : Self-Refining Radiology Report Generation Using Vision Language Models
Manav Nitin Kapadnis, Sohan Patnaik, Abhilash Nandy, Sourjyadip Ray, Pawan Goyal and Debdoot Sheet 283

ERD: A Framework for Improving LLM Reasoning for Cognitive Distortion Classification
Sehee Lim, Yejin Kim, Chi-Hyun Choi, Jy-yong Sohn and Byung-Hoon Kim 292

Leveraging Prompt-Learning for Structured Information Extraction from Crohn's Disease Radiology Reports in a Low-Resource Language
Liam Hazan, Naama Gavrielov, Roi Reichart, Talar Hagopian, Mary-Louise C. Greer, Ruth Cytter-Kuint, Gili Focht, Dan Turner and Moti Freiman 301

Context Aggregation with Topic-focused Summarization for Personalized Medical Dialogue Generation
Zhengyuan Liu, Siti Umairah Md Salleh, Pavitra Krishnaswamy and Nancy F. Chen 310

Evaluating Lexicon Incorporation for Depression Symptom Estimation
Kirill Milintsevich, Gaël Dias and Kairit Sirts 322

Semi-automatic Construction of a Word Complexity Lexicon for Japanese Medical Terminology
Soichiro Sugihara, Tomoyuki Kajiwara, Takashi Ninomiya, Shoko Wakamiya and Eiji Aramaki 329

TEAM MIPAL at MEDIQA-M3G 2024: Large VQA Models for Dermatological Diagnosis
Hyeonjin Kim, Min Kyu Kim, Jae Won Jang, KiYoon Yoo and Nojun Kwak 334

MediFact at MEDIQA-M3G 2024: Medical Question Answering in Dermatology with Multimodal Learning
Nadia Saeed . . . 339

MediFact at MEDIQA-CORR 2024: Why AI Needs a Human Touch
Nadia Saeed . . . 346

KnowLab_AIMed at MEDIQA-CORR 2024: Chain-of-Though (CoT) prompting strategies for medical error detection and correction
Zhaolong Wu, Abul Hasan, Jinge Wu, Yunsoo Kim, Jason Pui-Yin Cheung, Teng Zhang and Honghan Wu . . . 353

PromptMind Team at EHRSQL-2024: Improving Reliability of SQL Generation using Ensemble LLMs
Satya Kesav Gundabathula and Sriram R Kolar . . . 360

PromptMind Team at MEDIQA-CORR 2024: Improving Clinical Text Correction with Error Categorization and LLM Ensembles
Satya Kesav Gundabathula and Sriram R Kolar . . . 367

Maven at MEDIQA-CORR 2024: Leveraging RAG and Medical LLM for Error Detection and Correction in Medical Notes
Suramya Jadhav, Abhay Shanbhag, Sumedh Joshi, Atharva Date and Sheetal S. Sonawane . . . 374

LAILab at Chemotimelines 2024: Finetuning sequence-to-sequence language models for temporal relation extraction towards cancer patient undergoing chemotherapy treatment
Shohreh Haddadan, Tuan-Dung Le, Thanh Duong and Thanh Q. Thieu . . . 382

Lexicans at Chemotimelines 2024: Chemotimeline Chronicles - Leveraging Large Language Models (LLMs) for Temporal Relations Extraction in Oncological Electronic Health Records
Vishakha Sharma, Andres Fernandez, Andrei Constantin Ioanovici, David Talby and Frederik Buijs . . . 394

Team NLPeers at Chemotimelines 2024: Evaluation of two timeline extraction methods, can generative LLM do it all or is smaller model fine-tuning still relevant ?
Nesrine Bannour, Judith Jeyafreeda Andrew and Marc Vincent . . . 406

KCLab at Chemotimelines 2024: End-to-end system for chemotherapy timeline extraction – Subtask2
Yukun Tan, Merve Dede and Ken Chen . . . 417

Project PRIMUS at EHRSQL 2024 : Text-to-SQL Generation using Large Language Model for EHR Analysis
Sourav Bhowmik Joy, Rohan Ahmed, Argha Pratim Saha, Minhaj Ahmed Habil, Utsho Das and Partha Sarothi Bhowmik . . . 422

NYULangone at Chemotimelines 2024: Utilizing Open-Weights Large Language Models for Chemotherapy Event Extraction
Jeff Zhang, Yin Aphinyanaphongs and Anthony B Cardillo . . . 428

AIRI NLP Team at EHRSQL 2024 Shared Task: T5 and Logistic Regression to the Rescue
Oleg Somov, Alexey Dontsov and Elena Tutubalina . . . 431

IKIM at MEDIQA-M3G 2024: Multilingual Visual Question-Answering for Dermatology through VLM Fine-tuning and LLM Translations
Marie Bauer, Amin Dada, Constantin Marc Seibold and Jens Kleesiek . . . 439

NEUI at MEDIQA-M3G 2024: Medical VQA through consensus
Ricardo Omar Chávez García and Oscar William Lithgow-Serrano . . . 448

VerbaNexAI at MEDIQA-CORR: Efficacy of GRU with BioWordVec and ClinicalBERT in Error Correction in Clinical Notes
Juan Pajaro, Edwin Puertas, David Villate, Laura Estrada and Laura Tinjaca 461

HSE NLP Team at MEDIQA-CORR 2024 Task: In-Prompt Ensemble with Entities and Knowledge Graph for Medical Error Correction
Airat Valiev and Elena Tutubalina 470

Wonder at Chemotimelines 2024: MedTimeline: An End-to-End NLP System for Timeline Extraction from Clinical Narratives
Liwei Wang, Qiuhao Lu, Rui Li, Sunyang Fu and Hongfang Liu 483

Edinburgh Clinical NLP at MEDIQA-CORR 2024: Guiding Large Language Models with Hints
Aryo Pradipta Gema, Chaeeun Lee, Pasquale Minervini, Luke Daines, T. Ian Simpson and Beatrice Alex 488

UMass-BioNLP at MEDIQA-M3G 2024: DermPrompt - A Systematic Exploration of Prompt Engineering with GPT-4V for Dermatological Diagnosis
Parth Vashisht, Abhilasha Lodha, Mukta Maddipatla, Zonghai Yao, Avijit Mitra, Zhichao Yang, Sunjae Kwon, Junda Wang and Hong yu 502

KU-DMIS at MEDIQA-CORR 2024: Exploring the Reasoning Capabilities of Small Language Models in Medical Error Correction
Hyeon Hwang, Taewhoo Lee, Hyunjae Kim and Jaewoo Kang 526

CLD-MEC at MEDIQA- CORR 2024 Task: GPT-4 Multi-Stage Clinical Chain of Thought Prompting for Medical Errors Detection and Correction
Renad Alzghoul, Ayaabdelhaq Ayaabdelhaq, Abdulrahman Tabaza and Ahmad Altamimi . . . 537

Overview of the 2024 Shared Task on Chemotherapy Treatment Timeline Extraction
Jiarui Yao, Harry Hochheiser, WonJin Yoon, Eli T Goldner and Guergana K Savova 557

IryoNLP at MEDIQA-CORR 2024: Tackling the Medical Error Detection & Correction Task on the Shoulders of Medical Agents
Jean-Philippe Corbeil 570

Overview of the MEDIQA-M3G 2024 Shared Task on Multilingual Multimodal Medical Answer Generation
Wen-wai Yim, Asma Ben Abacha, Yujuan Fu, Zhaoyi Sun, Fei Xia, Meliha Yetisgen and Martin Krallinger 581

EM_Mixers at MEDIQA-CORR 2024: Knowledge-Enhanced Few-Shot In-Context Learning for Medical Error Detection and Correction
Swati Rajwal, Eugene Agichtein and Abeed Sarker 590

Overview of the MEDIQA-CORR 2024 Shared Task on Medical Error Detection and Correction
Asma Ben Abacha, Wen-wai Yim, Yujuan Fu, Zhaoyi Sun, Fei Xia and Meliha Yetisgen 596

UTSA-NLP at ChemoTimelines 2024: Evaluating Instruction-Tuned Language Models for Temporal Relation Extraction
Xingmeng Zhao and Anthony Rios 604

WangLab at MEDIQA-CORR 2024: Optimized LLM-based Programs for Medical Error Detection and Correction
Augustin Toma, Ronald Xie, Steven Palayew, Patrick Lawler and BO Wang 616

WangLab at MEDIQA-M3G 2024: Multimodal Medical Answer Generation using Large Language Models
Augustin Toma, Ronald Xie, Steven Palayew, Gary D. Bader and BO Wang 624

LG AI Research & KAIST at EHRSQL 2024: Self-Training Large Language Models with Pseudo-Labeled Unanswerable Questions for a Reliable Text-to-SQL System on EHRs
Yongrae Jo, Seongyun Lee, Minju Seo, Sung Ju Hwang and Moontae Lee 635

Overview of the EHRSQL 2024 Shared Task on Reliable Text-to-SQL Modeling on Electronic Health Records
Gyubok Lee, Sunjun Kweon, Seongsu Bae and Edward Choi 644

Saama Technologies at EHRSQL 2024: SQL Generation through Classification Answer Selector by LLM
Mohammed Jabir, Kamal Raj Kanakarajan and Malaikannan Sankarasubbu 655

KU-DMIS at EHRSQL 2024 : Generating SQL query via question templatization in EHR
Hajung Kim, Chanhwi Kim, Hoonick Lee, Kyochul Jang, Jiwoo Lee, Kyungjae Lee, Gangwoo Kim and Jaewoo Kang 672

ProbGate at EHRSQL 2024: Enhancing SQL Query Generation Accuracy through Probabilistic Threshold Filtering and Error Handling
Sangryul Kim, Donghee Han and Sehyun Kim 687

LTRC-IIITH at EHRSQL 2024: Enhancing Reliability of Text-to-SQL Systems through Abstention and Confidence Thresholding
Jerrin John Thomas, Pruthwik Mishra, Dipti Sharma and Parameswari Krishnamurthy 697

LTRC-IIITH at MEDIQA-M3G 2024: Medical Visual Question Answering with Vision-Language Models
Jerrin John Thomas, Sushvin Marimuthu and Parameswari Krishnamurthy 703

Program

Friday, June 21, 2024

09:00 - 09:10 *Opening Remarks*

09:10 - 09:50 *Keynote: David Restrepo*

09:50 - 10:00 *Keynote Q&A*

10:00 - 10:30 *Oral Session I*

Revisiting Clinical Outcome Prediction for MIMIC-IV
Tom Röhr, Alexei Figueroa, Jens-Michalis Papaioannou, Conor Fallon, Keno Bressem, Wolfgang Nejdl and Alexander Löser

A Multilevel Analysis of PubMed-only BERT-based Biomedical Models
Vicente Ivan Sanchez Carmona, Shanshan Jiang and Bin Dong

10:30 - 11:00 *Break*

11:00 - 11:30 *Oral Session II: Clinical Conversations*

Exploring Robustness in Doctor-Patient Conversation Summarization: An Analysis of Out-of-Domain SOAP Notes
Yu-Wen Chen and Julia Hirschberg

Can LLMs Correct Physicians, Yet? Investigating Effective Interaction Methods in the Medical Domain
Burcu Sayin, Pasquale Minervini, Jacopo Staiano and Andrea Passerini

11:30 - 12:00 *Oral Session III: Multimodal*

SERPENT-VLM : Self-Refining Radiology Report Generation Using Vision Language Models
Manav Nitin Kapadnis, Sohan Patnaik, Abhilash Nandy, Sourjyadip Ray, Pawan Goyal and Debdoot Sheet

Fusion of Domain-Adapted Vision and Language Models for Medical Visual Question Answering
Cuong Ha, Shima Asaadi, Sanjeev Kumar Karn, Oladimeji Farri, Tobias Heimann and Thomas Runkler

Friday, June 21, 2024 (continued)

12:00 - 12:30 *Oral Session IV: Low-Resource Settings*

A Privacy-Preserving Corpus for Occupational Health in Spanish: Evaluation for NER and Classification Tasks
Claudio Aracena, Luis Miranda, Thomas Vakili, Fabián Villena, Tamara Quiroga, Fredy Núñez-Torres, Victor Rocco and Jocelyn Dunstan

Leveraging Prompt-Learning for Structured Information Extraction from Crohn's Disease Radiology Reports in a Low-Resource Language
Liam Hazan, Naama Gavrielov, Roi Reichart, Talar Hagopian, Mary-Louise C. Greer, Ruth Cytter-Kuint, Gili Focht, Dan Turner and Moti Freiman

12:30 - 14:00 *Lunch*

14:00 - 14:22 *Oral Session V: MEDIQA-CORR*

Overview of the MEDIQA-CORR 2024 Shared Task on Medical Error Detection and Correction
Asma Ben Abacha, Wen-wai Yim, Yujuan Fu, Zhaoyi Sun, Fei Xia and Meliha Yetisgen

PromptMind Team at MEDIQA-CORR 2024: Improving Clinical Text Correction with Error Categorization and LLM Ensembles
Satya Kesav Gundabathula and Sriram R Kolar

14:22 - 14:45 *Oral Session VI: MEDIQA-M3G*

Overview of the MEDIQA-M3G 2024 Shared Task on Multilingual Multimodal Medical Answer Generation
Wen-wai Yim, Asma Ben Abacha, Yujuan Fu, Zhaoyi Sun, Fei Xia, Meliha Yetisgen and Martin Krallinger

WangLab at MEDIQA-M3G 2024: Multimodal Medical Answer Generation using Large Language Models
Augustin Toma, Ronald Xie, Steven Palayew, Gary D. Bader and BO Wang

14:45 - 15:07 *Oral Session VII: EHRSQL*

Overview of the EHRSQL 2024 Shared Task on Reliable Text-to-SQL Modeling on Electronic Health Records
Gyubok Lee, Sunjun Kweon, Seongsu Bae and Edward Choi

LG AI Research & KAIST at EHRSQL 2024: Self-Training Large Language Models with Pseudo-Labeled Unanswerable Questions for a Reliable Text-to-SQL System on EHRs
Yongrae Jo, Seongyun Lee, Minju Seo, Sung Ju Hwang and Moontae Lee

Friday, June 21, 2024 (continued)

15:07 - 15:30 *Oral Session VIII: ChemoTimelines*

Overview of the 2024 Shared Task on Chemotherapy Treatment Timeline Extraction
Jiarui Yao, Harry Hochheiser, WonJin Yoon, Eli T Goldner and Guergana K Savova

LAILab at Chemotimelines 2024: Finetuning sequence-to-sequence language models for temporal relation extraction towards cancer patient undergoing chemotherapy treatment
Shohreh Haddadan, Tuan-Dung Le, Thanh Duong and Thanh Q. Thieu

15:30 - 16:00 *Break*

16:00 - 17:30 *Poster Session*