



LREC 2026

**The 8th Workshop on Clinical Natural Language
Processing (Clinical NLP) @ LREC 2026**

Workshop Proceedings

Editors

Asma Ben Abacha

Steven Bethard

Danielle Bitterman

Tristan Naumann

Kirk Roberts

16 May 2026

©ELRA Language Resources Association (ELRA), 2026
These proceedings are licensed under a Creative Commons Attribution-
NonCommercial 4.0 International License (CC BY-NC 4.0)

ISBN 978-2-493814-56-2

Preface

This volume contains papers from the 8th Workshop on Clinical Natural Language Processing (Clinical NLP 2026).

Clinical text is growing rapidly as electronic health records (EHRs) become pervasive. 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 to better understand patient state and care provided. The goal of this workshop is to establish a regular event that brings together researchers interested in developing state-of-the-art methods for the clinical domain. The focus is on improving NLP technology to enable clinical applications via information extraction and modeling of narrative provider notes from electronic health records, patient encounter transcripts, and other clinical narratives. Clinical NLP 2026 also hosted two shared tasks, challenging researchers around the world to develop new approaches to solve clinical NLP problems: scoring model-generated answers to patient questions and extracting/normalizing clinical observations from conversational transcripts.

The main workshop received a total of 40 submissions, of which 7 were accepted as oral presentations, and 14 were accepted as poster presentations. The shared tasks received a total of 18 participant submissions, of which 2 were accepted as oral presentations, and 16 were accepted as poster presentations.

Keynote Talk

Holistic engagement beyond the edges of data-driven science: The case of Libya

Dr. Stephen Wu

Saraya Hamra University

Abstract: We often make some assumptions about the ethics (e.g., individual rights) and structures (e.g., research enterprise) supporting health technologies. This talk argues that, for societies like Libya's, some paradigm shifts are in order. First, we contrast ethical values between honor vs. dignity cultures. Second, we highlight aspects of the Libyan research system that fundamentally change what work needs to be done in health technology research. Subsequently, we describe several recent holistic, contextually situated efforts at engagement: private research sponsorship and administration, science of science research, and care practice norms and information flow in medicine.

Biography: Dr. Stephen Wu is an Associate Professor at Saraya Hamra University (SHU) in Tripoli, Libya, where he serves as the Director for the Center for Research & Innovation as well as the Head of the Department of Computer Science & Engineering. His current interests lie in empowering the Libyan innovation system through the science of science, the interaction of global value systems with modern AI, and culturally situated healthcare information systems.

Organizing Committee

General Chairs

Asma Ben Abacha, Microsoft, USA

Steven Bethard, University of Arizona, USA

Danielle Bitterman, Harvard University, USA

Tristan Naumann, Microsoft Research, USA

Kirk Roberts, University of Texas Health Science Center at Houston, USA

MEDIQA-EVAL 2026 Shared Task Organizers

Asma Ben Abacha, Microsoft, USA

Wen-wai Yim, Microsoft, USA

MEDIQA-SYNUR 2026 Shared Task Organizers

Asma Ben Abacha, Microsoft, USA

Nate Bodenstab, Microsoft, USA

Jean-Philippe Corbeil, Microsoft, Canada

George Michalopoulos, Microsoft, Canada

Reviewers

Abdulrahman AAIAbdulsalam, Bayu Aryoyudanta, Federica Bologna, Leonardo Campillos-Llanos, Jean-Philippe Corbeil, Yadan Fan, Haritha Gangavarapu, Natalia Grabar, Ziqi Hao, Harry Hochheiser, Sy Hwang, Kyomin Hwang, Yoshinobu Kano, Halil Kilicoglu, Primož Kocbek, Yanis Labrak, Egoitz Laparra, Alberto Lavelli, Ulf Leser, Yuxing Lu, Diwakar Mahajan, Sushvin Marimuthu, George Michalopoulos, Timothy Miller, Prateek Munjal, Aurélie Névél, Yifan Peng, Xinzhe Peng, Lisa Raithel, Satyajeet Raje, Pavithra Rajendran, Giridhar Kaushik Ramachandran, A H M Rezaul Karim, Frank Rudzicz, Michael Saban, Guergana Savova, Ankit Singh, Arvind Krishna Sridhar, Swetha Krishna Sriram, Lynda Tamine, Jerrin Thomas, Aashwin Vaish, Airat Valiev, V.G.Vinod Vydiswaran, Jinge Wu, Justin Xu, Dongfang Xu, Jiarui Yao, Wen-wai Yim, WonJin Yoon, Boya Zhang, Pierre Zweigenbaum

Table of Contents

<i>Overview of the MEDIQA-EVAL 2026 Shared Task on Evaluation Metrics in Medical Multimodal Question Answering</i>	
Asma Ben Abacha and Wen-wai Yim	1
<i>SUAT-BMI at MEDIQA-EVAL 2026: An Ensemble Approach to Language Models as Judges for Automatic Rating of Medical Responses</i>	
Xinzhe Peng, Liyuan E, Kun Feng, Jieli Li, Yuxuan Tang and Zhao Li	12
<i>Overview of the MEDIQA-SYNUR 2026 Shared Task on Observation Extraction from Nurse Dictations</i>	
George Michalopoulos, Jean-Philippe Corbeil, Cari Bader, Nathan Bodenshtab and Asma Ben Abacha	19
<i>SemAntICA Lab at MediQA-SYNUR 2026: Route, Extract and Verify – An LLM-gated Ensemble for Parsing Nurse Dictations</i>	
Sy Hwang, Katherine S. Pitcher, Sue Hyon Kim, Yoonjae Lee, Hayoung K. Donnelly, Harsh Bandhey, Andrew J. King, Karen O'Connor, Ryan J. Urbanowicz and Danielle L. Mowery ...	27
<i>L2D-Clinical: Learning to Defer for Adaptive Model Selection in Clinical Text Classification</i>	
Rishik Kondadadi and John E. Ortega	38
<i>TRUMEDIQA: A Modular Trustworthy RAG Pipeline for Multilingual Medical Question Answering</i>	
Meryem El Fatimi, Ayoub Nainia and Jihad Zahir	49
<i>Evaluating the Retrieval Component in a Retrieval-Augmented Summarization System for Patient Records in French</i>	
Marco Naguib, Christel Gérardin, Victor Beaucoté, Cyril Charron, Adrien Joseph, Aurélie Névél and Xavier Tannier	57
<i>Retrieval-Augmented Generation Based Nurse Observation Extraction</i>	
Kyomin Hwang and Nojun Kwak	66
<i>Automatic Generation of Discharge Summaries Using Large Language Models: A Systematic Literature Review</i>	
Lucas Molino-Piñar, Manuel Carlos Diaz Galiano and María-Teresa Martín-Valdivia	73
<i>Smart_solutions at MEDIQA-SYNUR 2026: A Multi-Stage LLM Pipeline for Nursing Observation Extraction</i>	
Prateek Munjal	83
<i>GS-BrainText: A Multi-Site Brain Imaging Report Dataset from Generation Scotland for Clinical Natural Language Processing Development and Validation</i>	
Beatrice Alex, Claire Grover, Arlene Casey, Richard Tobin, Heather Whalley and William Whiteley	92
<i>SASTA Self Assessment: An efficient human-in-the-loop strategy for developmental and pathological language analysis</i>	
Jan Odijk, Jelte van Boheemen, Xander Vertegaal, Tessel Boerma and Marijn Schraagen	103

<i>Differentially Private De-identification of Dutch Clinical Notes: A Comparative Evaluation</i> Michele Miranda, Xinlan Yan, Nishant Mishra, Rachel Murphy, Ameen Abu Hanna, Sébastien Bratières and Iacer Calixto	113
<i>SQUCS at MEDIQA-SYNUR 2026: A Multi-Agent Open Source LLM System for Nursing Observation Extraction</i> Riham JeebAllah, Adhari AlZaabi and Abdulrahman Khalifa AAIAbdulsalam	125
<i>LTRC-IIIT at MEDIQA-SYNUR 2026: Benchmarking a Fully Local, Training-Free RAG Pipeline</i> Aashwin Vaish and Dipti Misra Sharma	136
<i>BDI at MEDIQA-EVAL 2026: A ReAct-Style Multimodal Agent for Fine-Grained Medical Response Assessment</i> Justin Xu, Zizheng Zhang, Augustine Luk, Benjamin Khong, Haochen Cui, Samuel Hwang, Alyssa Pradhan, Kevin Yuan and David W. Eyre	141
<i>MasonNLP at MEDIQA-SYNUR 2026: Retrieval-Augmented Large Language Models for Schema-Constrained Clinical Information Extraction</i> A H M Rezaul Karim and Özlem Uzuner	153
<i>Night Shift Nerds at MEDIQA-SYNUR 2026: Pushing Small Large Language Model Capability for Clinical Observation Extraction and Normalization from Nurse Dictation using RLVR</i> Bayu Aryoyudanta, Maria Yuliana, Mikie Rachman and I Made Agus Setiawan	163
<i>Extracting Medication Instructions from Dutch General Practice Electronic Health Records with Local Natural Language Processing</i> Marya Dukmak, Constanza L. Andaur Navarro and Artuur Leeuwenberg	174
<i>Gladiator at MEDIQA-SYNUR 2026: Contextual Clinical Extraction: Integrating Foundation Models with Domain-Specific Validation Rules</i> Siva Satyanarayana Raju Pusapati and Ankit Singh	183
<i>MedAware at MEDIQA-EVAL 2026: Vision-Language Model Fine-Tuning with Logprob-Based Score Calibration for Medical Response Evaluation</i> Ziqi Hao and Pengbo Liu	192
<i>HSE NLP TEAM at MEDIQA-SYNUR 2026: Consensus Adjudication Ensemble (ACE): Balancing Precision and Recall for Schema-Bystander Clinical Extraction</i> Airat A. Valiev	200
<i>Lakefront AI Ramblers at MEDIQA-SYNUR 2026: Hybrid Retrieval and LLM Verification for Open-Source Schema-Guided Clinical Information Extraction</i> Michael T. Saban, Arsalan Yaghoubi, Behnaz Eslami, Samie Tootooni and Dmitriy Dligach	212
<i>JMedWiC: A Japanese Word-in-Context Dataset in the Medical Domain</i> Koki Horiguchi, Seiji Sugiyama, Tomoyuki Kajiwara, Shoko Wakamiya and Eiji Aramaki	222
<i>LTRC-Medicom at MEDIQA-SYNUR 2026: Schema-Guided Clinical Information Extraction with Hybrid Clustering-SFT-Verification</i> Pasumarthy Deepak, Sushvin Marimuthu and Parameswari Krishnamurthy	228

<i>SloCal-Net at MEDIQA-Eval 2026: Investigating the Impact of Reasoning and External Context on Medical Answer Grading</i>	
Primož Kocbek, Valentina Carbonari, Pierangelo Veltri, Pietro Hiram Guzzi and Gregor Stiglic	235
<i>MIDAS_SYNUR at MEDIQA-SYNUR 2026: A Prompting Study for Clinical Observation Extraction from Nurse Dictation Transcriptions</i>	
Swetha Krishna Sriram and Akshita Sahoo	244
<i>AnotherOne at MEDIQA-SYNUR 2026: Detect, Extract, Normalize - Knowledge-Grounded LLM Pipeline for Clinical Observation Extraction</i>	
Jerrin John Thomas and Parameswari Krishnamurthy	251
<i>hgkai26 at MEDIQA-EVAL 2026: Automated Evaluation of Visual Medical Question Answering Using LLM-as-a-Judge</i>	
Haritha Gangavarapu	257
<i>Role-Adapted Clinical Report Generation for Ultrasound Measurements in Low-Resource Settings</i>	
Ayoub Nainia, Tanya Akumu, Noussair Lazrak and Karim Lekadir	262
<i>A Comparative Study of Approaches to Anonymization of Clinical Free Text in Spanish</i>	
Florencia Luciana Brunello, Laura Alonso Alemany, Serena Villata and Milagro Teruel	270
<i>Disagreement-Driven Joint Refinement of Retrieval and Decision Rules for Imbalanced Counseling Risk Classification</i>	
Zhihao Shao, Ryo Sekizaki, Shengzhou Yi and Toshihiko Yamasaki	281
<i>Context-Aware SNOMED CT Entity Linking for Clinical Text</i>	
Provia Kadusabe, Demian Gholipour Ghalandari, Lauren Cassidy, Jack Boylan, Chris Hokamp, Abhishek Kaushik and Fiona Lawless	290
<i>Temporal Structure in Clinical Narratives in Portuguese: Insights from Cross-Document Annotation</i>	
Ana Luisa Fernandes, Purificação Silvano and Luís Filipe Cunha	300
<i>MOSAIC: A Multilingual, Taxonomy-Agnostic, and Computationally Efficient Approach for Radiological Report Classification in Low-Resource Settings</i>	
Alice Schiavone, Marco Fraccaro, Lea Marie Pehrson, Silvia Ingala, Rasmus Bonnevie, Michael Bachmann Nielsen, Vincent Beliveau, Melanie Ganz and Desmond Elliott	313
<i>MedNormJ: A Benchmark Dataset for Medical Concept Normalization in Japanese Clinical Documents</i>	
Yuki Tashiro, Seiji Shimizu, Tomohiro Nishiyama, Shoko Wakamiya and Eiji Aramaki	324
<i>Pediatric Sepsis Cohort Detection Using In-Context Pointwise V-Usable Information</i>	
Yingya Li, Alon Geva, Steven Bethard, Timothy A. Miller, Kate Madden, Matthew A. Eisenberg, Daniel P. Kelly and Guergana Savova	336
<i>RECAP: Transparent Inference-Time Emotion Alignment for Medical Dialogue Systems</i>	
Adarsh Srinivasan, Jacob Dineen, Muhammad Uzair Sarfraz, Muhammad Umar Afzal, Irbaz Riaz and Ben Zhou	350

<i>Disentangling Ambiguity from Instability in Large Language Models: A Clinical Text-to-SQL Case Study</i>	
Angelo Ziletti and Leonardo D'Ambrosi	369
<i>An OMOP-Based Open-Source Text-to-SQL Benchmark Dataset</i>	
Paul Legrand, Kawsar Noor, Satyam Bhagwanani and Richard J. Dobson	381
<i>Profiling Hallucinations in Frontier LLMs for Entity Linking to Medical Ontologies</i>	
Logan Born, Nishant Kambhatla, Uliyana Kubasova, Maryam Siahbani, Andrei Vacariu, Timothy W. O'Connell and Anoop Sarkar	394

Conference Program

Sat 16 May 2026

09:00–09:45 Opening

09:00–09:30 *Keynote: Holistic engagement beyond the edges of data-driven science: The case of Libya*
Dr. Stephen Wu

09:30–09:45 Keynote Q&A

09:45–10:00 MEDIQA-EVAL 2026 Shared Task

09:45–09:55 *Overview of the MEDIQA-EVAL 2026 Shared Task on Evaluation Metrics in Medical Multimodal Question Answering*
Asma Ben Abacha and Wen-wai Yim

09:55–10:00 *SUAT-BMI at MEDIQA-EVAL 2026: An Ensemble Approach to Language Models as Judges for Automatic Rating of Medical Responses*
Xinzhe Peng, Liyuan E, Kun Feng, Jieli Li, Yuxuan Tang and Zhao Li

10:00–10:15 MEDIQA-SYNUR 2026 Shared Task

10:00–10:10 *Overview of the MEDIQA-SYNUR 2026 Shared Task on Observation Extraction from Nurse Dictations*
George Michalopoulos, Jean-Philippe Corbeil, Cari Bader, Nathan Bodenstab and Asma Ben Abacha

10:10–10:15 *SemAnTICA Lab at MediQA-SYNUR 2026: Route, Extract and Verify – An LLM-gated Ensemble for Parsing Nurse Dictations*
Sy Hwang, Katherine S. Pitcher, Sue Hyon Kim, Yoonjae Lee, Hayoung K. Donnelly, Harsh Bandhey, Andrew J. King, Karen O'Connor, Ryan J. Urbanowicz and Danielle L. Mowery

10:15–11:15 Poster Session

10:15–11:15 *L2D-Clinical: Learning to Defer for Adaptive Model Selection in Clinical Text Classification*
Rishik Kondadadi and John E. Ortega

10:15–11:15 *TRUMEDIQA: A Modular Trustworthy RAG Pipeline for Multilingual Medical Question Answering*
Meryem El Fatimi, Ayoub Nainia and Jihad Zahir

Sat 16 May 2026 (continued)

- 10:15–11:15 *Evaluating the Retrieval Component in a Retrieval-Augmented Summarization System for Patient Records in French*
Marco Naguib, Christel Gérardin, Victor Beaucoté, Cyril Charron, Adrien Joseph, Aurélie Névéol and Xavier Tannier
- 10:15–11:15 *Retrieval-Augmented Generation Based Nurse Observation Extraction*
Kyomin Hwang and Nojun Kwak
- 10:15–11:15 *Automatic Generation of Discharge Summaries Using Large Language Models: A Systematic Literature Review*
Lucas Molino-Piñar, Manuel Carlos Diaz Galiano and María-Teresa Martín-Valdivia
- 10:15–11:15 *Smart_solutions at MEDIQA-SYNUR 2026: A Multi-Stage LLM Pipeline for Nursing Observation Extraction*
Prateek Munjal
- 10:15–11:15 *GS-BrainText: A Multi-Site Brain Imaging Report Dataset from Generation Scotland for Clinical Natural Language Processing Development and Validation*
Beatrice Alex, Claire Grover, Arlene Casey, Richard Tobin, Heather Whalley and William Whiteley
- 10:15–11:15 *SASTA Self Assessment: An efficient human-in-the-loop strategy for developmental and pathological language analysis*
Jan Odijk, Jelte van Boheemen, Xander Vertegaal, Tessel Boerma and Marijn Schraagen
- 10:15–11:15 *Differentially Private De-identification of Dutch Clinical Notes: A Comparative Evaluation*
Michele Miranda, Xinlan Yan, Nishant Mishra, Rachel Murphy, Ameen Abu Hanna, Sébastien Bratières and Iacer Calixto
- 10:15–11:15 *SQUCS at MEDIQA-SYNUR 2026: A Multi-Agent Open Source LLM System for Nursing Observation Extraction*
Riham JeebAllah, Adhari AlZaabi and Abdulrahman Khalifa AAIAbdul-salam
- 10:15–11:15 *LTRC-IIIT at MEDIQA-SYNUR 2026: Benchmarking a Fully Local, Training-Free RAG Pipeline*
Aashwin Vaish and Dipti Misra Sharma
- 10:15–11:15 *BDI at MEDIQA-EVAL 2026: A ReAct-Style Multimodal Agent for Fine-Grained Medical Response Assessment*
Justin Xu, Zizheng Zhang, Augustine Luk, Benjamin Khong, Haochen Cui, Samuel Hwang, Alyssa Pradhan, Kevin Yuan and David W. Eyre

Sat 16 May 2026 (continued)

- 10:15–11:15 *MasonNLP at MEDIQA-SYNUR 2026: Retrieval-Augmented Large Language Models for Schema-Constrained Clinical Information Extraction*
A H M Rezaul Karim and Özlem Uzuner
- 10:15–11:15 *Night Shift Nerds at MEDIQA-SYNUR 2026: Pushing Small Large Language Model Capability for Clinical Observation Extraction and Normalization from Nurse Dictation using RLVR*
Bayu Aryoyudanta, Maria Yuliana, Mikie Rachman and I Made Agus Setiawan
- 10:15–11:15 *Extracting Medication Instructions from Dutch General Practice Electronic Health Records with Local Natural Language Processing*
Marya Dukmak, Constanza L. Andaur Navarro and Artuur Leeuwenberg
- 10:15–11:15 *Gladiator at MEDIQA-SYNUR 2026: Contextual Clinical Extraction: Integrating Foundation Models with Domain-Specific Validation Rules*
Siva Satyanarayana Raju Pusapati and Ankit Singh
- 10:15–11:15 *MedAware at MEDIQA-EVAL 2026: Vision-Language Model Fine-Tuning with Logprob-Based Score Calibration for Medical Response Evaluation*
Ziqi Hao and Pengbo Liu
- 10:15–11:15 *HSE NLP TEAM at MEDIQA-SYNUR 2026: Consensus Adjudication Ensemble (ACE): Balancing Precision and Recall for Schema-Bystander Clinical Extraction*
Airat A. Valiev
- 10:15–11:15 *Lakefront AI Ramblers at MEDIQA-SYNUR 2026: Hybrid Retrieval and LLM Verification for Open-Source Schema-Guided Clinical Information Extraction*
Michael T. Saban, Arsalan Yaghoubi, Behnaz Eslami, Samie Tootooni and Dmitriy Dligach
- 10:15–11:15 *JMedWiC: A Japanese Word-in-Context Dataset in the Medical Domain*
Koki Horiguchi, Seiji Sugiyama, Tomoyuki Kajiwarra, Shoko Wakamiya and Eiji Aramaki
- 10:15–11:15 *LTRC-Medicom at MEDIQA-SYNUR 2026: Schema-Guided Clinical Information Extraction with Hybrid Clustering-SFT-Verification*
Pasumarthy Deepak, Sushvin Marimuthu and Parameswari Krishnamurthy
- 10:15–11:15 *SloCal-Net at MEDIQA-Eval 2026: Investigating the Impact of Reasoning and External Context on Medical Answer Grading*
Primož Kocbek, Valentina Carbonari, Pierangelo Veltri, Pietro Hiram Guzzi and Gregor Stiglic
- 10:15–11:15 *MIDAS_SYNUR at MEDIQA-SYNUR 2026: A Prompting Study for Clinical Observation Extraction from Nurse Dictation Transcriptions*
Swetha Krishna Sriram and Akshita Sahoo

Sat 16 May 2026 (continued)

- 10:15–11:15 *AnotherOne at MEDIQA-SYNUR 2026: Detect, Extract, Normalize - Knowledge-Grounded LLM Pipeline for Clinical Observation Extraction*
Jerrin John Thomas and Parameswari Krishnamurthy
- 10:15–11:15 *hgkai26 at MEDIQA-EVAL 2026: Automated Evaluation of Visual Medical Question Answering Using LLM-as-a-Judge*
Haritha Gangavarapu
- 10:15–11:15 *Role-Adapted Clinical Report Generation for Ultrasound Measurements in Low-Resource Settings*
Ayoub Nainia, Tanya Akumu, Noussair Lazrak and Karim Lekadir
- 10:15–11:15 *A Comparative Study of Approaches to Anonymization of Clinical Free Text in Spanish*
Florencia Luciana Brunello, Laura Alonso Alemany, Serena Villata and Milagro Teruel
- 10:15–11:15 *Disagreement-Driven Joint Refinement of Retrieval and Decision Rules for Imbalanced Counseling Risk Classification*
Zhihao Shao, Ryo Sekizaki, Shengzhou Yi and Toshihiko Yamasaki
- 10:15–11:15 *Context-Aware SNOMED CT Entity Linking for Clinical Text*
Provia Kadusabe, Demian Gholipour Ghalandari, Lauren Cassidy, Jack Boylan, Chris Hokamp, Abhishek Kaushik and Fiona Lawless
- 10:15–11:15 *Temporal Structure in Clinical Narratives in Portuguese: Insights from Cross-Document Annotation*
Ana Luisa Fernandes, Purificação Silvano and Luís Filipe Cunha
- 11:15–11:45 **Special Track on Low Resource Settings**
- 11:15–11:30 *MOSAIC: A Multilingual, Taxonomy-Agnostic, and Computationally Efficient Approach for Radiological Report Classification in Low-Resource Settings*
Alice Schiavone, Marco Fraccaro, Lea Marie Pehrson, Silvia Ingala, Rasmus Bonnevie, Michael Bachmann Nielsen, Vincent Beliveau, Melanie Ganz and Desmond Elliott
- 11:30–11:45 *MedNormJ: A Benchmark Dataset for Medical Concept Normalization in Japanese Clinical Documents*
Yuki Tashiro, Seiji Shimizu, Tomohiro Nishiyama, Shoko Wakamiya and Eiji Aramaki

Sat 16 May 2026 (continued)

**11:45–
12:15** **Special Track on Construction or Deconstruction of LLMs**

11:45–
12:00 *Pediatric Sepsis Cohort Detection Using In-Context Pointwise V-Usable Information*

Yingya Li, Alon Geva, Steven Bethard, Timothy A. Miller, Kate Madden, Matthew A. Eisenberg, Daniel P. Kelly and Guergana Savova

12:00–
12:15 *RECAP: Transparent Inference-Time Emotion Alignment for Medical Dialogue Systems*

Adarsh Srinivasan, Jacob Dineen, Muhammad Uzair Sarfraz, Muhammad Umar Afzal, Irbaz Riaz and Ben Zhou

**12:15–
13:00** **Main Track**

12:15–
12:30 *Disentangling Ambiguity from Instability in Large Language Models: A Clinical Text-to-SQL Case Study*

Angelo Ziletti and Leonardo D'Ambrosi

12:30–
12:45 *An OMOP-Based Open-Source Text-to-SQL Benchmark Dataset*

Paul Legrand, Kawsar Noor, Satyam Bhagwanani and Richard J. Dobson

12:45–
13:00 *Profiling Hallucinations in Frontier LLMs for Entity Linking to Medical Ontologies*

Logan Born, Nishant Kambhatla, Uliyana Kubasova, Maryam Siahbani, Andrei Vacariu, Timothy W. O'Connell and Anoop Sarkar

13:00 **Closing**

