

SMM4H 2026

**The 11th Social Media Mining for Health Research and
Applications (SMM4H-HeaRD 2026) Workshop and Shared
Tasks**

Proceedings of the Workshop

July 3, 2026

©2026 Association for Computational Linguistics

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 979-8-89176-432-3

Preface

Welcome to the 11th Social Media Mining for Health Research and Applications and the Health Real-World Data (#SMM4H-HeaRD) Workshop and Shared Tasks, co-located with the 64th Annual Meeting of the Association for Computational Linguistics (ACL 2026). This year’s workshop is held online, connecting the global community of researchers who develop and evaluate natural language processing, machine learning, and artificial intelligence methods for health-related text.

Since its inception in 2016, #SMM4H has grown from a small, focused workshop, into one of the field’s most visible recurring venues for health NLP research. Over eleven editions, the workshop has brought together hundreds of teams from across the world, spanning academia, industry, government, and clinical institutions. Participation has grown substantially year over year. Collectively, the #SMM4H shared tasks and associated publications have accumulated over 1,700 citations, reflecting the community’s sustained reliance on the datasets, baselines, and benchmarking frameworks the workshop has produced.

The #SMM4H-HeaRD 2026 edition marks a record in the scope of our shared task program, with eight shared tasks spanning an unprecedentedly broad range of domains and data modalities: detection of adverse drug events in multilingual social media posts, detection of insomnia in clinical notes, estimating flu vaccine effectiveness from social media, generation of structured medical notes from dialogues, detection of patient metadata in SARS-CoV-2 sequencing articles, predicting TNM staging from pathology reports, extraction of social and clinical impacts of substance use, and multilingual clinical entity annotation projection and extraction. This year, 110 teams from 31 countries registered. The shared tasks included 79 task-level team participations across the 8 tasks: 12 teams for Task 1, 8 for Task 2, 6 for Task 3, 5 for Task 4, 10 for Task 5, 7 for Task 6, 10 for Task 7, and 21 for Task 8, the largest participation in the workshop’s history. Winning systems were invited for oral presentations, and other accepted system description papers were included as poster presentations. In all, these proceedings include 53 papers.

This year’s edition also highlights the field’s evolving methodological landscape. Large language models featured prominently across submissions, used for classification, information extraction, data augmentation, and structured output generation - reflecting both their growing accessibility and the community’s interest in evaluating their reliability in high-stakes health contexts. In alignment with the ACL 2026 Theme Track on the Explainability of NLP Models, we specifically invited work that advances the transparency, interpretability, and trustworthiness of models applied in clinical and public health settings.

We are deeply grateful to our program committee, reviewers, and, particularly our shared task organizers for their rigorous and dedicated work. We also thank the annotators who created the shared task datasets, the ACL 2026 organizing team for their support, and every researcher who submitted a paper or participated in a shared task. #SMM4H-HeaRD would not exist without this community, and we are grateful to see it continue to grow.

Graciela Gonzalez Hernandez, PhD - Cedars-Sinai Health Sciences University, General Chair
Guillermo López-García, PhD - Stanford University, 2026 Co-Chair
Ari Klein, PhD - University of Pennsylvania, 2026 Co-Chair

Organizing Committee

General Chair

Graciela Gonzalez-Hernandez, Cedars-Sinai Medical Center, USA

Workshop Co-Chairs

Guillermo Lopez-Garcia, Stanford Health Care, USA

Ari Klein, University of Pennsylvania, USA

Shared Task Organizers

Graciela Gonzalez-Hernandez, Cedars-Sinai Medical Center, USA

Guillermo Lopez-Garcia, Stanford Health Care, USA

Dongfang Xu, Independent Researcher, Former Research Data Scientist at Cedars-Sinai Medical Center, CA, USA

Ivan Flores, Cedars-Sinai Medical Center, USA

Jose Miguel Acitores, Cedars-Sinai Medical Center, USA

Jacob Berkowitz, Cedars-Sinai Medical Center, USA

Nicholas Tatonetti, Cedars-Sinai Medical Center, USA

Ari Z. Klein, University of Pennsylvania, USA

Abeed Sarker, Emory University, USA

Sumon Kanti Dey, Emory University, USA

Ahmad Rezaie Mianroodi, Dalhousie University, Vector Institute, Canada

Frank Rudzicz, Dalhousie University, Vector Institute, Canada

Lisa Raithel, Technische Universität Berlin, Germany

Roland Roller, DFKI, Germany

Philippe Thomas, DFKI, Germany

Farnaz Zeidi, Paul Ehrlich Institute, Germany

Yu Zhai, Hong Kong Polytechnic University, Hong Kong

Tomohiro Nishiyama, Nara Institute of Science and Technology, Japan

Pierre Zweigenbaum, LISN, CNRS, Université Paris-Saclay, France

Elena Tutbalina, AIRI, Kazan Federal University, Russia

Salvador Lima-López, Barcelona Supercomputing Center, Spain

Judith Rosell, Barcelona Supercomputing Center, Spain

Martin Krallinger, Barcelona Supercomputing Center, Spain

Fernando Gallego-Donoso, Barcelona Supercomputing Center, Spain

Workshop Committee

Graciela Gonzalez-Hernandez, Cedars-Sinai Medical Center, USA

Guillermo Lopez-Garcia, Stanford Health Care, USA

Ari Klein, University of Pennsylvania, USA

Yujun Ma, Cedars-Sinai Medical Center, USA

Publication Chair

Guillermo Lopez-Garcia, Stanford Health Care, USA

Program Committee

Program Chairs

Lisa Raithel
Guillermo Lopez-Garcia
Dongfang Xu
Ahmad Rezaie Mianroodi
Ari Klein
Jacob Berkowitz
Jose Miguel Acitores
Sumon Kanti Dey
Abeed Sarker
Salvador Lima-López
Fernando Gallego-Donoso

Program Committee

Juan M Banda
Robert Leaman
Luis M Rocha
Kirk Roberts
Raul Rodriguez-Esteban
Elena Tutubalina
Karin Verspoor
Pierre Zweigenbaum

Reviewers

Radja Afren
Rathijit Aich
João Almeida
Ioan-Tudor-Alexandru Anghel
Sophie Arnoult
Vasudev Awatramani
Parsa Bagherzadeh
Juan Banda
Shankar Biradar
Tiberiu Boros
Svetla Boytcheva
Diego Burgos
Suchandra Chakraborty
Donger Chen
Ivan Chernov
Radu-Gabriel Chivereanu
Emilia-Ioana Cristea
Gijs Danoe
Nirjhar Das
Harshal Dharpure

Andres Duque
Diego Estuar
Rubén Fernández
Pietro Ferrazzi
Dhruv Goyal
Ishita Gupta
Brian Habersberger
Antonio Jesus Tamayo Herrera
Yi-Wen Hsiao
Sofia Jamil
Richard Jonker A. A.
Ayush Kumar
Shobhan Kumar
Alberto Lavelli
Robert Leaman
Zihan Liang
Abhishek Maity
Alexandros Maniatis
Sérgio Matos
Francisco Moreno-Barea
Abir Naskar
Abdusalam Nwesri
Karen O'Connor
Bhaarat Pachori
Vasile Pais
Irina Patularu
Rajesh Piryani
Aatish Pradhan
Supreetha R
François Remy
Kirk Roberts
Luis Rocha
Raul Rodriguez-Esteban
Thanya Santhosh
Yutaka Sasaki
Hsuan-Lei Shao
Aman Sinha
Soubraylu Sivakumar
Zafi Sherhan Syed
Ramya Tekumalla
Doan Tien
Nicolas Turenne
Elena Tutubalina
Evita Vardhani
Sylvia Vassileva
Francisco Veredas
Karin Verspoor
Piek Vossen
Thin Vãn
Deahan Yu
Farnaz Zeidi

Petr Zelina
Huixin Zhan
Pierre Zweigenbaum
Bram van Es

Keynote Talk

Beyond Model Performance: What Kind of Care Does AI Enable?

Wendy Chapman
UT Southwestern Medical Center

Abstract: As AI becomes increasingly capable of performing tasks once reserved for clinicians, the most important questions are shifting from algorithms to systems. This talk will examine how AI is changing the nature of clinical work, patient participation, and healthcare delivery, and why model accuracy alone tells us little about real-world impact. Drawing on experiences from Learning Health Systems, digital health implementation, and human-centered design, I will argue that evaluation must begin long before deployment and continue throughout the innovation lifecycle. Ultimately, the future of clinical NLP depends on moving beyond asking whether AI can perform a task and instead ask a more fundamental question: What kind of care does AI enable?

Bio: Wendy Chapman, PhD, is Associate Dean of Health Informatics and Chief Learning Health Officer at UT Southwestern Medical Center. She got a BA in linguistics and followed her husband into the field of medical informatics, and for more than 20 years, she developed and evaluated natural language processing and AI methods for healthcare, with the goal of improving patient care. Over time, she realized that strong model performance alone rarely translated into changes in clinical practice, leading her to focus on implementation, human factors, and learning health systems. After helping establish the Centre for Digital Transformation of Health at the University of Melbourne to support the translation of digital innovations into care, she now works at the intersection of research and health system operations, bridging the gap between innovation and implementation.

Keynote Talk

“I’ll Be Your Mirror”: Using NLP to Expand Patient Meaning Making

Brian Chapman
UT Southwestern Medical Center

Abstract: NLP healthcare applications have almost exclusively been clinician facing, even though patients have had increasing access, by law and by practice, to their own clinical texts. A systematic review of NLP on patient-authored text found that even work using patient-generated data has been oriented primarily toward clinical surveillance and institutional quality improvement; the patient who generated the text rarely benefits directly from the analysis. Yet the tools themselves are not inherently clinician-facing or applicable only to clinical text. This talk explores how NLP can serve as a mirror through which patients see themselves reflected in their own data. Drawing on a fifty-year longitudinal auto-ethnographic corpus blending memoir, personal journals, medical records, wearable data, and other artifacts, I juxtapose what clinical records document with what personal narratives reveal, illustrating meaning making across diverse health circumstances using tools ranging from psycholinguistic text analysis to generative AI. I close by suggesting that NLP-enabled reflections can meaningfully support patient meaning making, and by sketching open questions for patient-facing NLP research

Bio: Brian Chapman is an associate professor at UT Southwestern Medical Center, with an honorary appointment at the University of Melbourne. His academic journey began in electrical engineering and applied mathematics before the siren call of health research - dormant since his childhood as a cancer patient - lured him away from a plasma physics PhD at the University of Wisconsin–Madison toward medical imaging. He completed his PhD at the University of Utah, where his dissertation tackled medical imaging problems from a medical informatics perspective.

For much of his career Brian’s research was primarily within radiology, until a move to Australia in 2019 - necessitated, as with most of his many institutional moves, by his wife’s globetrotting habits - brought a different health system, a different university culture, and a global pandemic that together redirected his interests toward using informatics and AI to help patients better engage with their own healthcare. He views his greatest professional accomplishments as having occurred in the classroom.

Outside of work, Brian vacuums after three huskies, makes visual art, and reads widely in biography, memoir, history, and philosophy. A native of Salt Lake City, he loves downhill skiing and mountain biking. His favorite TV shows include Yes, Minister, The Rockford Files, and Fawlty Towers, which he considers a perfectly coherent set of preferences.

Table of Contents

<i>A3S@C-DAC at #SMM4H-HeaRD 2026: Reasoning Meets Evidence: LLMs for Interpretable Insomnia Detection with Evidence Extraction in Clinical Notes</i>	
Abhishek Maity, Amol Shinde, Abhishek Suresh Kushare and Swapnil Pawar	1
<i>Gladiators at #SMM4H-HeaRD 2026: Multi-Seed XLM-RoBERTa Ensemble with Focal Loss and Per-Language Threshold Optimization for Multilingual Adverse Drug Event Detection</i>	
Ankit Kumar Singh	7
<i>LSI_UNED at #SMM4H-HeaRD 2026: Grid-Based Biomedical Named Entity Recognition Across Languages and Entity Types</i>	
Alicia Ramirez-Arrabe, Juan Martinez-Romo and Andres Duque	12
<i>SINAI at #SMM4H-HeaRD 2026: Multilingual Clinical NER with MrBERT-biomed and Optuna Hyperparameter Optimization</i>	
Lucas Molino Piñar, Manuel Carlos Diaz-Galiano and María-Teresa Martín-Valdivia	18
<i>Prestige at #SMM4H-HeaRD 2026: Binary Insomnia Classification from Clinical Notes Using LLMs with Chain-of-Thought Reasoning</i>	
Oyindolapo O. Komolafe	23
<i>Team Gazoo! at #SMM4H-HeaRD 2026: Zero-Training NER via Iterative LLM Prompt Self-Optimization for Opioid Impact Span Detection</i>	
Diego Estuar	28
<i>DNT at #SMM4H-HeaRD 2026: Leveraging BERT-based Encoders and LLMs for Medical Information Extraction</i>	
Doan Nhat Tien and Thìn Đăng Văn	36
<i>BIT.UA at #SMM4H-HeaRD 2026: Towards Multi-Class Multilingual Clinical Entity Recognition with Multi-Head CRF Ensembles</i>	
Richard A. A. Jonker and Sérgio Matos	41
<i>Bhramastra at #SMM4H-HeaRD 2026: A Multi-Stage Hunter-Judge Pipeline using DSPy-Optimized LLMs for Multilingual ADE Detection</i>	
Bhaarat Pachori	49
<i>LLATMU at #SMM4H-HeaRD 2026: Clinical Text Structuring with QLoRA-based Generation and Partial-Label TNM Classification</i>	
Eric Hsiao, Min-Hsuan Ku and Hsuan-Lei Shao	56
<i>Patient2Paper at #SMM4H-HeaRD 2026: Retrieval-Augmented Few-Shot Generation for Clinical Note Synthesis</i>	
Ioan-Tudor-Alexandru Anghel, Timotei Andrei, Comârdici Marian Bogdan and Carina Sâicu	61
<i>In2Lab-TNT at #SMM4H-HeaRD 2026: An Application of QTT's Terminological Entanglement to Leverage Insomnia Detection in Clinical Notes</i>	
Antonio Jesus Tamayo Herrera, Giovanni Díaz-Laínes, Carlos Mario Perez Perez and Diego A Burgos	67
<i>blue at SMM4H-HeaRD 2026: Class-Weighted Transformer Ensembles with Structured Decoding and Chain-of-Thought Blending across Six Health NLP Shared Tasks</i>	
Krish Sharma, Rhea Singhal and Jatin Bedi	72

<i>DT4H.nl at #SMM4H-HeaRD 2026: Multilingual Clinical NER with multilingual and monolingual models</i>	
Bram van Es	82
<i>SMMTech at #SMM4H-HeaRD 2026: Detection of Insomnia in Clinical Notes</i>	
Emilia-Ioana Cristea	88
<i>Beyond Lexical Similarity: Evaluating Faithfulness in LLM-Based Medical Question Reformulation</i>	
Md Rabiul Hasan, Aleka Melese Ayalew and Mourad Oussalah	93
<i>NU_DeepHealthNLP at #SMM4H-HeaRD 2026: Entity-Conditioned Generation and a Four-Stage Pipeline for Automated SOAP Note Generation</i>	
Thanya Mysore Santhosh and Deahan Yu	103
<i>GoBlueInformatics at #SMM4H-HeaRD 2026: Long-Context Encoders and Generative Biomedical LLMs for Pathological TNM Stage Prediction</i>	
Shangqing Wei	108
<i>Enigma at #SMM4H-HeaRD 2026: Leveraging Multilingual Pre-trained Models for Clinical Named Entity Recognition</i>	
Sylvia Vassileva, Plamena Ilieva, Teodor Svetoslavov Kostadinov, Monika Peteva Petkova, Daniel Manchevski, Vitosh Doynov, Ivan Koychev and Svetla Boytcheva	113
<i>RACAI at #SMM4H-HeaRD: Named Entity Recognition for Detecting the Impacts of Drug Abuse in Social Media Posts: Zero-Shot and Fine-Tuning Approaches</i>	
Tiberiu Boros and Radu-Gabriel Chivoreanu	121
<i>ICB-UMA at #SMM4H-HeaRD 2026: Hybrid Clinical Entity Projection for MultiClinAI: Adaptive Candidate Windows, XGBoost, and LLM Refinement</i>	
Alvaro Rey-Blanes, Sara Giménez-Gómez, Francisco J. Veredas and Francisco J. Moreno-Barea	127
<i>URJC-Team at #SMM4H-HeaRD 2026: TNM Stage Extraction with a Regex-LLM Workflow</i>	
Natalia Madrueño, Jose Walter Hernández Pérez, Rubén R. Fernández and Soto Montalvo ...	133
<i>LotusOrchid at #SMM4H-HeaRD 2026: Fitting pretrained encoders for Dutch medical data</i>	
Sophie Arnoult, Shutao Chen and Piek Vossen	139
<i>PEI at #SMM4H-HeaRD 2026: Enhancing Patient Metadata Detection via Hypothesis-Conditioned Classification and Paraphrase-Based Data Augmentation</i>	
Farnaz Zeidi, Roman Christof, Farnoush Zeidi, Renate König and Liam Childs	146
<i>Dr-BERT-NL at #SMM4H-HeaRD 2026: DOKTERBERT – Ontology-Grounded Contextual Representations for Dutch Clinical NLP</i>	
Gijs Danoe, Andreas Voss, Axel Hamprecht and Matthijs S. Berends	154
<i>Vasudev Awatramani at #SMM4H-HeaRD 2026: A Two-Pass LLM Pipeline with Deterministic Rule Derivation for Interpretable Insomnia Detection in Clinical Notes</i>	
Vasudev Awatramani	160
<i>Parallia at #SMM4H-HeaRD 2026: ClinicalAligner26AM: A Cross-Lingual Aligner for Dataset Translation; Evidences from the MultiClinCorpus Shared Task</i>	
François Remy	165
<i>Discovery@FI at #SMM4H-HeaRD 2026: Ensemble Character Classifier for Multilingual Clinical NER</i>	
Petr Zelina and Vit Novacek	173

<i>IITPatna_ADE at #SMM4H-HeaRD 2026: Multilingual Adverse Drug Event Detection with LoRA-XLM-RoBERTa, Cross-Fold Ensembles, and Post-hoc Calibration</i>	
Sofia Jamil, Manish Singh, Harshal Dharpure, Sriparna Saha and Rajiv Misra	177
<i>CUET_DiagNLP at #SMM4H-HeaRD 2026: Per-Axis TNM Staging from Pathology Reports and Opioid Impact Span Detection from Social Media</i>	
Shuva Dey, Priyangshu Barua and Mohammad Ashfak Habib	182
<i>MedMind AI at #SMM4H-HeaRD 2026: Data Extraction and Generation Using Prompt Engineering and Structured Outputs (Tasks 1–6)</i>	
Aatish Pradhan and Brian M. Habersberger	187
<i>CaresAI at SMM4H-HeaRD 2026: Predicting TNM Staging</i>	
Joseph Itopa Abubakar, Jorge Jarne, Favour Igwezeke and Mary Adewunmi	206
<i>Vinland_Vector at #SMM4H-HeaRD 2026: Multilingual ADE Detection and Query-Augmented Clinical NER for English</i>	
Nirjhar Das, Rathijit Aich and Mahfuzulhoq Chowdhury	211
<i>SIEMENS at #SMM4H-HeaRD 2026: The Impact of Training Strategy and Backbone Selection on BERT-based Multilingual Clinical NER</i>	
Manuela Daniela Danu	216
<i>HALELab-NITK at #SMM4H-HeaRD2026: Inclusion of Feature Engineering for Detection of Patient Metadata in SARS-CoV2 Sequencing Articles</i>	
Aakarsh Bansal, Abhishek Srinivas and Sowmya Kamath S.	222
<i>Cuet_Data_Wizards at #SMM4H-HeaRD 2026: Multilingual ADE Detection and Influenza Vaccine Effectiveness Estimation from Social Media</i>	
Abir Dey, Mohammed Omar Faiaz and Muhammad Ibrahim Khan	225
<i>Limics at #SMM4H-HeaRD 2026: Uncertainty-Driven Prediction for ADE Detection in Social Media</i>	
Nour Allam	230
<i>FU-HU-P5 at #SMM4H-HeaRD 2026: MedSynth Dialogue-to-Note Generation</i>	
Jessica Ying En Wong	237
<i>ACSS-PSL at #SMM4H-HeaRD 2026: An LLM-Driven Autoresearch Loop for Opioid-Impact NER</i>	
Olivier Caron, Bruno Chaves Ferreira and Christophe Benavent	240
<i>Creative Catalysts at #SMM4H-HeaRD 2026: XLM-RoBERTa for Task 1 Binary Classification of Social Media Posts Containing Adverse Drug Events</i>	
Radja Afren, Hichem Rahab and Imane Guellil	246
<i>BioNLP at #SMM4H-HeaRD 2026 Task 3 Estimating Flu Vaccine Effectiveness: A Temporal-Aware Fine-Tuning and Similarity-Based Few-Shot Prompting Approach</i>	
Irina Patularu	252
<i>Infimobius at #SMM4H-HeaRD 2026: Multi-Seed DeBERTa Ensemble for Flu Vaccination and Testing Status Classification</i>	
Pradyumn Kejriwal, Suhani Singh Charan, Raksha Sharma and Rudra Murthy	260
<i>Thunderbolts at #SMM4H-HeaRD 2026: Detection of Insomnia in Clinical Notes using Transformers</i>	
Guddanti Venkata Sree Charan, nama_ss@cs.iitr.ac.in, nama_ss@cs.iitr.ac.in, Raksha Sharma and Rudra Murthy	264

<i>Team TIET at #SMM4H-HeaRD 2026: Fine-tuned Biomedical Transformers with Language-Balanced Sampling for Patient Metadata and Multilingual ADE Detection</i>	
Divrose Kaur, Jatin Bedi and Jasmeet Singh	268
<i>MetaMiners at SMM4H-HeaRD 2026: A Semantic-Structural Knowledge-Enriched Ensemble for SARS-CoV-2 Metadata Identification</i>	
Claudia-Alexandra Ursu and Alecsandru-Florin Soare	272
<i>No_gmail at #SMM4H-HeaRD 2026: Detecting Patient Metadata in COVID-19 Scientific Literature: A Comparative Study of Encoder-Only and Autoregressive Language Models</i>	
Stefanescu Anastasia	280
<i>Understanding the Sociocultural Dimensions of Mental Health Discourse in Arabic X Communities</i>	
Amal Abdullah Alqahtani, Rana Aref Salama and Mona T. Diab	286
<i>Team Paradise at #SMM4H-HeaRD 2026: Multi-Task Approaches for Social Media Health Mining</i>	
Dhruv Goyal, Ishita Gupta and Jatin Bedi	305
<i>The MultiClinAI Shared Task on Multilingual Clinical Corpus Construction and Concept Extraction: Systems, Evaluation, and Datasets</i>	
Fernando Gallego Donoso, Salvador Lima-Lopez, Judith Rosell, Eulàlia Farré-Maduel and Martin Krallinger	309
<i>Overview of #SMM4H-HeaRD 2026 – Task 6: Predicting TNM staging from pathology reports</i>	
Jose Miguel Acitores Cortina, Jacob S. Berkowitz, Nadine A. Friedrich and Nicholas P Tatonetti	332
<i>NoviceTrio in #SMM4H-HeaRD 2026: Hybrid Clinical Transformer Ensembles for Insomnia Detection and Evidence Extraction from Clinical Notes</i>	
Abir Naskar and Mike Conway	338
<i>Overview of #SMM4H-HeaRD 2026 - Task 2: Detection of Insomnia in Clinical Notes</i>	
Joey Chan, Lauren D. Gryboski, Guillermo Lopez-Garcia and Graciela Gonzalez-Hernandez	345
<i>Overview of the 11th Social Media Mining for Health (#SMM4H) and Health Real-World Data (HeaRD) Shared Tasks at ACL 2026</i>	
Guillermo Lopez-Garcia, Jose Miguel Acitores Cortina, Jacob Berkowitz, Joey Chan, Sumon Kanti Dey, Ivan Flores Amaro, Fernando Gallego, Lauren Gryboski, Ari Z. Klein, Farnoush Zeidi Kolehparcheh, Martin Krallinger, Salvador Lima-Lopez, Yujun Ma, Tomohiro Nishiyama, Ahmad Rezaie Mianroodi, Amirali Rezaie Mianroodi, Lisa Raithel, Roland Roller, Judith Rosell, Frank Rudzicz, Abeed Sarker, Nicholas Tatonetti, Philippe Thomas, Elena Tutubalina, Dongfang Xu, Farnaz Zeidi, Yu Zhai, Pierre Zweigenbaum and Graciela Gonzalez-Hernandez	353

Program

Friday, July 3, 2026

- 08:30 - 08:45 *Workshop Introduction – Graciela Gonzalez-Hernandez*
- 08:45 - 09:20 *Task 8 - MultiClinAI: Multilingual Clinical Entity Annotation Projection and Extraction*
- BIT.UA at #SMM4H-HeaRD 2026: Towards Multi-Class Multilingual Clinical Entity Recognition with Multi-Head CRF Ensembles*
Richard A. A. Jonker and Sérgio Matos
- Parallia at #SMM4H-HeaRD 2026: ClinicalAligner26AM: A Cross-Lingual Aligner for Dataset Translation; Evidences from the MultiClinCorpus Shared Task*
François Remy
- 09:20 - 09:45 *Task 1 - Detection of Adverse Drug Events in Multilingual and Multi-platform Social Media Posts*
- Bhramastra at #SMM4H-HeaRD 2026: A Multi-Stage Hunter-Judge Pipeline using DSPy-Optimized LLMs for Multilingual ADE Detection*
Bhaarat Pachori
- 09:45 - 10:00 *Workshop Paper*
- Understanding the Sociocultural Dimensions of Mental Health Discourse in Arabic X Communities*
Amal Abdullah Alqahtani, Rana Aref Salama and Mona T. Diab
- 10:00 - 10:25 *Task 4 - Generation of Realistic Structured Medical Notes from Dialogues*
- NU_DeepHealthNLP at #SMM4H-HeaRD 2026: Entity-Conditioned Generation and a Four-Stage Pipeline for Automated SOAP Note Generation*
Thanya Mysore Santhosh and Deahan Yu
- 10:25 - 11:00 *Coffee Break*
- 11:00 - 11:25 *Task 5 - Detection of Patient Metadata in SARSCoV-2 Sequencing Articles*
- MetaMiners at SMM4H-HeaRD 2026: A Semantic-Structural Knowledge-Enriched Ensemble for SARS-CoV-2 Metadata Identification*
Claudia-Alexandra Ursu and Alecsandru-Florin Soare

Friday, July 3, 2026 (continued)

11:25 - 12:15 *Keynote Speaker 1 – Wendy Chapman*

12:15 - 13:15 *Lunch Break*

13:15 - 13:40 *Task 6 - Predicting TNM Staging from TCGA Pathology Reports*

URJC-Team at #SMM4H-HeaRD 2026: TNM Stage Extraction with a Regex-LLM Workflow

Natalia Madrueño, Jose Walter Hernández Pérez, Rubén R. Fernández and Soto Montalvo

13:40 - 13:55 *Task 3 - Estimating Flu Vaccine Effectiveness from Social Media Posts*

BioNLP at #SMM4H-HeaRD 2026 Task 3 Estimating Flu Vaccine Effectiveness: A Temporal-Aware Fine-Tuning and Similarity-Based Few-Shot Prompting Approach

Irina Patularu

13:55 - 14:40 *Keynote Speaker 2 – Brian Chapman*

14:40 - 15:40 *Poster Presentation Session*

15:40 - 15:55 *Workshop Paper*

Beyond Lexical Similarity: Evaluating Faithfulness in LLM-Based Medical Question Reformulation

Md Rabiul Hasan, Aleka Melese Ayalew and Mourad Oussalah

15:55 - 16:20 *Task 7 - Extraction of Social and Clinical Impacts of Substance Use from Social Media Posts*

Team Gazoo! at #SMM4H-HeaRD 2026: Zero-Training NER via Iterative LLM Prompt Self-Optimization for Opioid Impact Span Detection

Diego Estuar

16:20 - 16:45 *Task 2 - Detection of Insomnia in Clinical Notes*

MedMind AI at #SMM4H-HeaRD 2026: Data Extraction and Generation Using Prompt Engineering and Structured Outputs (Tasks 1–6)

Aatish Pradhan and Brian M. Habersberger

Friday, July 3, 2026 (continued)

16:45 - 17:00 *Conclusion and Closing Remarks – Graciela Gonzalez-Hernandez*