

ACL 2026

**BioNLP 2026 - Shared Tasks**

**Proceedings of the 25th Workshop on Biomedical Language  
Processing (Shared Tasks)**

July 3-4, 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](mailto:acl@aclweb.org)

ISBN 979-8-89176-435-4

# Introduction

*Deepak Gupta and Dina Demner-Fushman*

The BioNLP workshop has a long-standing tradition of offering shared tasks to advance biomedical natural language processing through thorough, novel, and timely topics, evaluation, and community engagement. These shared tasks foster the development of robust and reproducible methods, evaluation frameworks, and datasets, as well as clinically relevant research challenges for the community. This year, we invited SIGBioMed members to submit a description of a shared task proposal. We reviewed the proposals and selected four shared tasks to include in BioNLP 2026. The selected tasks cover a diverse and timely set of research directions, including detecting and labeling medical decisions in ICU discharge summaries (MedExACT), detecting and classifying psychological defense mechanisms in emotional support dialogues (PsyDefDetect), multimodal biomedical question answering with grounded attribution (MedGenVidQA), and clinical skill understanding and continuous perception from clinical video frames (ClinicalSkillQA). Overall, these shared tasks foster the development and benchmarking of innovative methods for multimodal information retrieval, information extraction, question answering, clinical psychology, and mental health in biomedicine.

A total of 39 papers were submitted across the four shared tasks, with participants detailing a range of novel approaches and systems. We accepted 35 papers for inclusion in the proceedings. Each shared task is briefly described below.

**MedExACT** The first MedExACT shared task focused on detecting and labeling medical decisions in ICU discharge summaries. A medical decision is defined by the Decision Identification and Classification Taxonomy for Use in Medicine (DICTUM), which covers ten categories, including Defining Problem, Therapeutic Procedure, and Evaluation. The MedDec dataset contains over 56k expert-annotated decision spans drawn from de-identified ICU discharge summaries in MIMIC-III, supplemented with patient demographics and ten phenotypes (Cancer, Heart Disease, Lung Disease, Chronic Neurologic Dystrophies, Chronic Pain, Alcohol Abuse, Substance Abuse, Obesity, Psychiatric Disorders, Depression). Systems were evaluated for accuracy and robustness at both span and token levels, including stratified analyses by sex, race, English proficiency, and disease type. Baseline models such as RoBERTa indicated the complexity of the task, and participants were supported through expedited access to MedDec via PhysioNet, a public leaderboard, and a Python starter kit.

**PsyDefDetect** PsyDefDetect invited researchers to tackle a novel challenge at the intersection of Clinical Psychology and Natural Language Processing: detecting and classifying psychological defense mechanisms in emotional support dialogues. Grounded in the clinically validated Defense Mechanism Rating Scales (DMRS) framework, this shared task aimed to advance the understanding of unconscious defensive functioning in text. The benchmark addressed the challenge of capturing subtle linguistic cues of deep-seated psychological mechanisms within highly informal and context-dependent emotional dialogues. This initiative supported research at the intersection of clinical psychology and NLP, aiming to operationalize complex psychological constructs for computational analysis. Participating systems were ranked using Accuracy, Precision, Recall, and F1-score.

**MedGenVidQA** The recent surge in the availability of online videos has changed the way information and knowledge are acquired. Many people prefer instructional videos to learn how to accomplish a particular task effectively and efficiently through a series of step-by-step procedures. This need is not limited only to general audiences. In professional settings such as healthcare, instructional videos are widely used by physicians and other professionals to learn, review, and standardize procedural workflows. In addition, consumers increasingly seek step-by-step visual explanations to better understand medical procedures and clinical practices. With the advancement of generative models, the medical domain

has also seen progress in medical video understanding, clinical decision support, etc. Toward this, the MedGenVidQA shared task focused on developing systems that utilized generative models to retrieve relevant multimodal (textual and visual) sources and to localize visual answers within medical videos in response to consumer and healthcare professional medical queries. Following earlier editions of medical question answering tasks—MedVidQA 2023, MedVidQA 2024, BioGen 2024, and BioGen 2025—this shared task expanded medical video question answering for both professionals and consumers, with a focus on generative approaches to solving these tasks.

**ClinicalSkillQA** Multimodal large language models (MLLMs) have the potential to support clinical training and assessment by assisting medical experts in interpreting procedural videos and verifying adherence to standardized workflows. Reliable deployment in these settings requires evidence that models can continuously interpret students’ actions during clinical skill assessments, which underpins MLLMs’ understanding of clinical skills. Systematically evaluating and improving MLLMs’ understanding of clinical skills and their continuous perception in clinical skill assessment scenarios is therefore essential for building reliable and high-impact AI systems for medical education. To address this need, the shared task on medical question answering targeted clinical skill assessment scenarios.

## Organizing Committee

### MedExACT

Mohamed Elgaar, University of Massachusetts Lowell, USA  
Jiali Cheng, University of Massachusetts Lowell, USA  
Nidhi Vakil, University of Massachusetts Lowell, USA  
Mehrnaz Sadrolashrafi, Beth Israel Deaconess Medical Center, USA  
Mitra Mohtarami, Saint Anselm College, USA  
Adrian Wong, Beth Israel Deaconess Medical Center, USA  
Hadi Amiri, University of Massachusetts Lowell, USA  
Leo A. Celi, Massachusetts Institute of Technology, USA

### PsyDefDetect

Hongbin Na, University of Technology Sydney, Australia  
Zimu Wang, Xi'an Jiaotong-Liverpool University, China  
Zhaoming Chen, University of Utah, USA  
Yining Hua, Harvard University, USA  
Rena Gao, The University of Melbourne, Australia  
Kailai Yang, The University of Manchester, United Kingdom  
Ling Chen, University of Technology Sydney, Australia  
Wei Wang, Xi'an Jiaotong-Liverpool University, China  
Shaoxiong Ji, ELLIS Institute Finland & University of Turku, Finland  
John Torous, Harvard University, USA  
Sophia Ananiadou, The University of Manchester, United Kingdom

### MedGenVidQA

Deepak Gupta, National Library of Medicine, National Institutes of Health (NIH), USA  
Dina Demner-Fushman, National Library of Medicine, National Institutes of Health (NIH), USA

### ClinicalSkillQA

Xiyang Huang, Wuhan University, China  
Yihuai Xu, Wuhan University, China  
Zhiyuan Chen, Wuhan University, China  
Keying Wu, Wuhan University, China  
Jiayi Xiang, Wuhan University, China  
Buzhou Tang, Harbin Institute of Technology, China  
Renxiong Wei, Zhongnan Hospital of Wuhan University, China  
Yanqing Ye, Zhongnan Hospital of Wuhan University, China  
Jinyu Chen, Zhongnan Hospital of Wuhan University, China  
Cheng Zeng, Wuhan University, China  
Min Peng, Wuhan University, China  
Qianqian Xie, Wuhan University, China  
Sophia Ananiadou, The University of Manchester, United Kingdom

# Program Committee

## Chairs

Deepak Gupta, National Library of Medicine, NIH  
Dina Demner-Fushman, National Library of Medicine, NIH

## Reviewers

Hadi Amiri, University of Massachusetts Lowell  
Jiali Cheng, Northeastern University  
Mohamed Elgaar, University of Massachusetts Lowell  
Deepak Gupta, National Library of Medicine, NIH  
Yining Hua, Harvard University  
Xiyang Huang, Wuhan University  
Nidhi Vakil, University of Massachusetts Lowell  
Hilmi Demirhan, University of North Carolina Wilmington  
Gulustan Dogan, University of North Carolina Wilmington  
Basil Ebinesar, Purdue University Northwest  
Pedram Golnari, Case Western Reserve University  
Watcharitpol Sermsrisuwan, Department of Biomedical Engineering, Faculty of Engineering, Mahidol University  
Hongbin Na, University of Technology Sydney

## Table of Contents

<i>NM at CRF Filling 2026: A Two-Stage LLM Pipeline for Clinical CRF Population</i> Niccolò Morabito .....	1
<i>VerbaNexAI at ClinicalSkillQA: From Visual Evidence to Procedural Order A Two-Stage Generative Vision-Language Framework for ClinSkillQA</i> Andrea Menco Tovar, Jairo E. Serrano, Edwin Puertas and Juan Carlos Martinez-Santos.....	6
<i>zzucs at PsyDefDetect: Bridging Long-Tail Imbalance and Clinical Rubrics for DMRS Defense-Level Detection</i> Bin Huang, Liuyuan Su, Kaixuan Yuan, Guanghui Zhao, Shixin Zhang and Kunli Zhang .....	13
<i>zzunlp at ClinicalSkillQA: Perceive-and-Plan with Decomposed In-Context Learning and Saliency-Guided Perception for Clinical Skill Keyframe Reordering</i> Bin Huang, Yi Luo, Zhontian Hua, Guanghui Zhao, Kaixuan Yuan and Kunli Zhang.....	24
<i>DLNLP at ClinicalSkillQA: EvidenceFlow for Structured Zero-Shot Clinical Keyframe Ordering</i> Kexin Li, Zhekun Wang, Yiran Wang and Di Zhao.....	33
<i>UTS at PsyDefDetect: Multi-Agent Councils and Absence-Based Reasoning for Defense Mechanism Classification</i> Dima Galat and Marian Rizoiu .....	38
<i>Otter at MedExAct2026: Diverse Encoder Ensemble for Medical Decision Span Detection</i> Lalita Lowphansirikul and Piyalitt Ittichaiwong .....	47
<i>Eraserhead at PsyDefDetect: Prompt Design and Class Rebalancing for Psychological Defense Mechanism Detection</i> Muhammad Abu Horaira, Mehreen Rahman and Nahian Chowdhury .....	54
<i>Nürnberg NLP at PsyDefDetect: Multi-Axis Voter Ensembles for Psychological Defence Mechanism Classification</i> Philipp Steigerwald, Eric Rudolph and Jens Albrecht .....	59
<i>Neural Nexus at PsyDefDetect: Fine-Tuning RoBERTa with Focal Loss and Role-Tagged Dialogue History for Defense Level Detection</i> Subhrajyoti Basu .....	66
<i>ELiRF-UPV@MedExACT 2026: Dynamic Section Conditioning for Medical Decision Span Detection in Discharge Summaries</i> Vicent Ahuir, Lluís Hurtado and María Castro-Bleda .....	71
<i>VISHC at PsyDefDetect: Mitigating Data Scarcity in Psychological Defense Classification with Context-Aware Synthetic Augmentation</i> Hoang-Thuy-Duong Vu, Quoc-Cuong Pham and Huy-Hieu Pham .....	77
<i>Diverse Transformer Ensemble with Majority Voting for Medical Decision Extraction at MedExACT 2026</i> Rishik Kondadadi .....	87
<i>FBK-NLP at ClinSkill QA 2026: Improving Temporal Reasoning via Keypoint-Augmented Inputs</i> Pedro Gabriel Campana, Alberto Lavelli and Bernardo Magnini .....	92

<i>transformer_1376 at PsyDefDetect: A QLoRA-Based Generative Framework for Context-Aware Psychological Defense Mechanism Detection</i>	
Pritha Saha, Shuvodwip Saha and Anik Mahmud Shanto	99
<i>Explainers at PsyDefDetect: Hierarchical Prompting and Representation-Based Classification for Psychological Defenses</i>	
Liudmila Babakova, Christopher Luongo-Vazquez and Ilia Stepin	104
<i>PerceptionLab at PsyDefDetect: Overcoming Extreme Response Bias in LLMs via Rubric-Grounded Retrieval and Supervised Clinical Reasoning Distillation for Fine-Grained Ordinal Classification</i>	
Tamjid Fahim, Syed Johan and Saad Bin Maksud	109
<i>LinguUTics at PsyDefDetect: Iterative Imbalance-Aware Fine-tuning of Qwen3-8B for Psychological Defense Mechanism Classification</i>	
Shefayat Adib, Ahmed Sani, Md Hasibur Alif and Ajwad Abrar	123
<i>TONI-NLP at PsyDefDetect: Defense Mechanism Detection via LLM-based Ensemble Methods</i>	
Durjoy Paul, Arshitha Basavaraj, Callum Chan, Veronica Perez-Rosas, Diana Inkpen, Francisco Pereira and Juan Antonio Lossio-Ventura	132
<i>Zero-Shot, Fine-Tuned, and Retrieval-Augmented Extraction of Clinical Decisions with Corpus Boundary Diagnostics</i>	
Mohammed Alliheedi, Robert Mercer, Anemily Machina, Sudipta Roy, Yetian Wang and Xindi Wang	141
<i>CASPAR: A Context-Aware Span Refinement Approach for Decision Support</i>	
Jing Tao, Amir Eskandari and Farhana Zulkernine	146
<i>KCL-Cogstack at PsyDefDetect: A Hierarchical Approach to Detecting Defense Mechanisms in Supportive Dialogue</i>	
Shubham Agarwal, Thomas Searle and Richard Dobson	155
<i>DAL Team at PsyDefDetect: From Supervised Encoders to Hierarchical LLM-RAG for Psychological Defense Detection</i>	
Anh Chu, Luong Tran, Dat Do, Phuong Mai, Quynh Le and Cat Can	164
<i>CUAMC @ MedExACT 2026: Robust Ensemble Voting for Fair Medical Decision Extraction</i>	
William Baumgartner and Lisa Schilling	171
<i>LAMAR at MedExACT 2026: Agreement-Driven Large Language Model Ensembles for Clinical Decision Extraction from Discharge Summaries</i>	
Monrada Chiewhawan, Keetawan Limaroon and Titipat Achakulvisut	179
<i>CS_Metro at PsyDefDetect: Detecting Psychological Defense Mechanisms in Mental Health Dialogues with Summarization-Enhanced Transformer Ensembles</i>	
Oarisa Rebayet, Radiul Walee, Symom Hossain Shohan, Kawsar Ahmed and Mohammed Moshiul Hoque	191
<i>Sparse Category Routing and Fairness-Aware Optimization for Medical Decision Extraction</i>	
Ahmed Elshehaby, Mohamed Abdalla and Youssef Mohamed	201
<i>AlienAnnotators at PsyDefDetect: What Lies Between the Lines: Probing Lightweight Open-Source LLMs for Psychological Defense Mechanism Detection</i>	
Siam Karip and Nahid Hossain	213

<i>Team Aurum at MedExACT 2026@ACL: Data Augmentation and Clinical Longformer Fine-Tuning for Medical Decision Extraction</i>	
Jyoti Kumari, Vinay Ulli and Anindita Mondal . . . . .	224
<i>NJUST-KMG at MedGenVidQA 2026: Cascade Multi-modal Alignment with Gaussian Soft Priors for Medical Visual Answer Localization</i>	
Jinglong Li and Yang Yang . . . . .	229
<i>LAMAR-2 at MedGenVidQA 2026: Visual Answer Localization in Medical Videos via Multimodal LLM and Context-Augmented Prompting</i>	
Watcharitpol Sermsrisuwan, Nopporn Lekuthai, Seksan Yoadsanit and Titipat Achakulvisut .	233
<i>Varja-Dominators at MedGenVidQA 2026: Hybrid Video and Document Retrieval using PubMedBERT, T5 Query Expansion, and Cross-Encoder Re-Ranking</i>	
Pratik Dhaktode, Suhani Bighane and Anupama Phakatkar . . . . .	243
<i>Pride-Boiler at MedGenVidQA 2026: LLM-Augmented BM25 Retrieval with Corrective Self-Verification for Biomedical Evidence Retrieval</i>	
Basil Ebinesar, Keyuan Jiang, Charansai Maddineni and Ashok Raja . . . . .	248
<i>Seahawk at MedGenVidQA 2026: LLM Segment-Range Selection for Medical Visual Answer Localization</i>	
Xiaotian Tian and Gulustan Dogan . . . . .	257
<i>UNCC at MedGenVidQA 2026: Structured Temporal Grounding for Medical Video Question Answering</i>	
Hilmi Demirhan and Wlodek Zadrozny . . . . .	262

# Program

## Saturday, July 4, 2026

09:00 - 09:10 *Opening remarks*

09:10 - 09:20 *MedExACT Overview recapitulation*

09:20 - 09:50 *Shared Task Papers*

*CUAMC @ MedExACT 2026: Robust Ensemble Voting for Fair Medical Decision Extraction*

William Baumgartner and Lisa Schilling

*Otter at MedExAct2026: Diverse Encoder Ensemble for Medical Decision Span Detection*

Lalita Lowphansirikul and Piyalitt Ittichaiwong

*LAMAR at MedExACT 2026: Agreement-Driven Large Language Model Ensembles for Clinical Decision Extraction from Discharge Summaries*

Monrada Chiewhawan, Keetawan Limaroon and Titipat Achakulvisut

09:50 - 10:00 *Talk by Moustafa Yehia Hassan — The Divergence Hypothesis: Unmasking Lexical Interference and Label Bias in Mental Health NLP*

10:00 - 10:10 *Talk by Sanya Bathla Taneja — CNET: Context Engineering Framework for Normalization of Clinical Trial Procedures*

10:10 - 10:20 *Talk by Niloofar Arazkhani — KALIMBA: Knowledge-Assisted Literature Mining for Biological Interaction Analysis*

10:20 - 10:30 *Talk by Krushil Bhojani — BioRAG: A Systematic Ablation Study of Retrieval Strategies for Biomedical Question Answering*

10:30 - 11:00 *Coffee Break*

11:00 - 11:10 *MedGenVidQA Overview recap*

*Varja-Dominators at MedGenVidQA 2026: Hybrid Video and Document Retrieval using PubMedBERT, T5 Query Expansion, and Cross-Encoder Re-Ranking*

Pratik Dhaktode, Suhani Bighane and Anupama Phakatkar

*Pride-Boiler at MedGenVidQA 2026: LLM-Augmented BM25 Retrieval with Corrective Self-Verification for Biomedical Evidence Retrieval*

Basil Ebinesar, Keyuan Jiang, Charansai Maddineni and Ashok Raja

**Saturday, July 4, 2026 (continued)**

*Seahawk at MedGenVidQA 2026: LLM Segment-Range Selection for Medical Visual Answer Localization*

Xiaotian Tian and Gulustan Dogan

*UNCC at MedGenVidQA 2026: Structured Temporal Grounding for Medical Video Question Answering*

Hilmi Demirhan and Wlodek Zadrozny

11:50 - 12:00 *Talk by Kadir Bulut Ozler — Analyzing Prompt Design Choices in Biomedical Information Extraction for Low-Resource Languages*

12:00 - 12:10 *Talk by Santiago Martínez and Lina Gomez — MeSHClass-ES and AnatEM-ES: Open Resources for Spanish Biomedical NLP*

12:10 - 12:20 *Talk by Lana Yeganova — BioTopicXplor: A Web Tool for Interactive Exploration of PubMed Literature through Reproducible Topics*

12:20 - 12:30 *Talk by Won Kim — Learning to Combine AI Annotations for Improved Biomedical Relevance Labeling*

12:30 - 14:00 *Lunch*

14:00 - 14:10 *PsyDefDetect Overview recap*

14:10 - 14:40 *Shared Task Papers*

*UTS at PsyDefDetect: Multi-Agent Councils and Absence-Based Reasoning for Defense Mechanism Classification*

Dima Galat and Marian RizoIU

*KCL-Cogstack at PsyDefDetect: A Hierarchical Approach to Detecting Defense Mechanisms in Supportive Dialogue*

Shubham Agarwal, Thomas Searle and Richard Dobson

*Nürnberg NLP at PsyDefDetect: Multi-Axis Voter Ensembles for Psychological Defence Mechanism Classification*

Philipp Steigerwald, Eric Rudolph and Jens Albrecht

14:40 - 14:50 *Talk by Olga Pelloni — Evaluation of Multilingual Text Simplification for the Mental Health Domain: Exploring Small Language Models*

14:50 - 15:00 *Talk by Vishwaa Shah — A Multi-View Framework for Cross-Domain Nutrition Misinformation Detection in Social Media*

**Saturday, July 4, 2026 (continued)**

- 15:00 - 15:10 *Talk by Vishwaa Shah — Small LLMs for Biomedical Claim Verification: Cost-Effective Fine-Tuning, Structural Dataset Shortcuts, and Cross-Domain Generalization*
- 15:10 - 15:20 *Talk by Francois Remy — Diagnosable ColBERT: Debugging Late-Interaction Retrieval Models Using a Learned Latent Space as Reference*
- 15:20 - 15:30 *Talk by Hui Wang — CrossDDI: Cross-Source Evidence-Grounded Drug-Drug Interaction Verification*
- 15:30 - 16:00 *Coffee Break*
- 16:00 - 16:10 *ClinicalSkillQA Overview recap*
- 16:10 - 16:30 *Shared Task Papers*
- zzunlp at ClinicalSkillQA: Perceive-and-Plan with Decomposed In-Context Learning and Saliency-Guided Perception for Clinical Skill Keyframe Reordering*  
Bin Huang, Yi Luo, Zhontian Hua, Guanghui Zhao, Kaixuan Yuan and Kunli Zhang
- FBK-NLP at ClinSkill QA 2026: Improving Temporal Reasoning via Keypoint-Augmented Inputs*  
Pedro Gabriel Campana, Alberto Lavelli and Bernardo Magnini
- 16:30 - 16:40 *Talk by Yixian Ma — VERICITE: Evaluating Sentence-Level Citation Faithfulness in Retrieval-Augmented Medical Question Answering*
- 16:40 - 16:50 *Talk by Rez Samantha Floresca — Forgotten Words: Benchmarking NeoBERT for Dementia Detection in Low-Resource Conversational Filipino and English Speech*
- 16:50 - 17:00 *Talk by Brandon Colelough — Towards Grounded Hallucination Definitions for Biomedical Question Answering with Reproducible Examples*
- 17:00 - 18:00 *Poster presentation by Mahule Roy — Towards Unified Factuality Evaluation for Biomedical QA and Summarization: Aligning Metrics with Clinical Use-Cases*
- 17:00 - 18:00 *Poster presentation by Yikun Han — When Evidence Conflicts: Uncertainty and Order Effects in Retrieval-Augmented Biomedical Question Answering*
- 17:00 - 18:00 *Shared Task Posters*
- VerbaNexAI at ClinicalSkillQA: From Visual Evidence to Procedural Order A Two-Stage Generative Vision-Language Framework for ClinSkillQA*  
Andrea Menco Tovar, Jairo E. Serrano, Edwin Puertas and Juan Carlos Martinez-Santos

Saturday, July 4, 2026 (continued)

*zzucs at PsyDefDetect: Bridging Long-Tail Imbalance and Clinical Rubrics for DMRS Defense-Level Detection*

Bin Huang, Liuyuan Su, Kaixuan Yuan, Guanghui Zhao, Shixin Zhang and Kunli Zhang

*zzunlp at ClinicalSkillQA: Perceive-and-Plan with Decomposed In-Context Learning and Saliency-Guided Perception for Clinical Skill Keyframe Reordering*

Bin Huang, Yi Luo, Zhontian Hua, Guanghui Zhao, Kaixuan Yuan and Kunli Zhang

*DLNLP at ClinicalSkillQA: EvidenceFlow for Structured Zero-Shot Clinical Keyframe Ordering*

Kexin Li, Zhekun Wang, Yiran Wang and Di Zhao

*UTS at PsyDefDetect: Multi-Agent Councils and Absence-Based Reasoning for Defense Mechanism Classification*

Dima Galat and Marian Rizoiu

*Otter at MedExAct2026: Diverse Encoder Ensemble for Medical Decision Span Detection*

Lalita Lowphansirikul and Piyalitt Ittichaiwong

*Eraserhead at PsyDefDetect: Prompt Design and Class Rebalancing for Psychological Defense Mechanism Detection*

Muhammad Abu Horaira, Mehreen Rahman and Nahian Chowdhury

*Nürnberg NLP at PsyDefDetect: Multi-Axis Voter Ensembles for Psychological Defence Mechanism Classification*

Philipp Steigerwald, Eric Rudolph and Jens Albrecht

*Neural Nexus at PsyDefDetect: Fine-Tuning RoBERTa with Focal Loss and Role-Tagged Dialogue History for Defense Level Detection*

Subhrajyoti Basu

*ELiRF-UPV@MedExACT 2026: Dynamic Section Conditioning for Medical Decision Span Detection in Discharge Summaries*

Vicent Ahuir, Lluís Hurtado and María Castro-Bleda

*VISHC at PsyDefDetect: Mitigating Data Scarcity in Psychological Defense Classification with Context-Aware Synthetic Augmentation*

Hoang-Thuy-Duong Vu, Quoc-Cuong Pham and Huy-Hieu Pham

*Diverse Transformer Ensemble with Majority Voting for Medical Decision Extraction at MedExACT 2026*

Rishik Kondadadi

**Saturday, July 4, 2026 (continued)**

*FBK-NLP at ClinSkill QA 2026: Improving Temporal Reasoning via Keypoint-Augmented Inputs*

Pedro Gabriel Campana, Alberto Lavelli and Bernardo Magnini

*transformer\_1376 at PsyDefDetect: A QLoRA-Based Generative Framework for Context-Aware Psychological Defense Mechanism Detection*

Pritha Saha, Shuvodwip Saha and Anik Mahmud Shanto

*Explainers at PsyDefDetect: Hierarchical Prompting and Representation-Based Classification for Psychological Defenses*

Liudmila Babakova, Christopher Luongo-Vazquez and Ilia Stepin

*PerceptionLab at PsyDefDetect: Overcoming Extreme Response Bias in LLMs via Rubric-Grounded Retrieval and Supervised Clinical Reasoning Distillation for Fine-Grained Ordinal Classification*

Tamjid Fahim, Syed Johan and Saad Bin Maksud

*LinguUTics at PsyDefDetect: Iterative Imbalance-Aware Fine-tuning of Qwen3-8B for Psychological Defense Mechanism Classification*

Shefayat Adib, Ahmed Sani, Md Hasibur Alif and Ajwad Abrar

*TONI-NLP at PsyDefDetect: Defense Mechanism Detection via LLM-based Ensemble Methods*

Durjoy Paul, Arshitha Basavaraj, Callum Chan, Veronica Perez-Rosas, Diana In-ken, Francisco Pereira and Juan Antonio Lossio-Ventura

*Zero-Shot, Fine-Tuned, and Retrieval-Augmented Extraction of Clinical Decisions with Corpus Boundary Diagnostics*

Mohammed Alliheedi, Robert Mercer, Anemily Machina, Sudipta Roy, Yetian Wang and Xindi Wang

*CASPAR: A Context-Aware Span Refinement Approach for Decision Support*

Jing Tao, Amir Eskandari and Farhana Zulkernine

*KCL-Cogstack at PsyDefDetect: A Hierarchical Approach to Detecting Defense Mechanisms in Supportive Dialogue*

Shubham Agarwal, Thomas Searle and Richard Dobson

*DAL Team at PsyDefDetect: From Supervised Encoders to Hierarchical LLM-RAG for Psychological Defense Detection*

Anh Chu, Luong Tran, Dat Do, Phuong Mai, Quynh Le and Cat Can

*CUAMC @ MedExACT 2026: Robust Ensemble Voting for Fair Medical Decision Extraction*

William Baumgartner and Lisa Schilling

**Saturday, July 4, 2026 (continued)**

*LAMAR at MedExACT 2026: Agreement-Driven Large Language Model Ensembles for Clinical Decision Extraction from Discharge Summaries*

Monrada Chiewhawan, Keetawan Limaroon and Titipat Achakulvisut

*CS\_Metro at PsyDefDetect: Detecting Psychological Defense Mechanisms in Mental Health Dialogues with Summarization-Enhanced Transformer Ensembles*

Oarisa Rebayet, Radiul Walee, Symom Hossain Shohan, Kawsar Ahmed and Mohammed Moshiul Hoque

*Sparse Category Routing and Fairness-Aware Optimization for Medical Decision Extraction*

Ahmed Elshehaby, Mohamed Abdalla and Youssef Mohamed

*AlienAnnotators at PsyDefDetect: What Lies Between the Lines: Probing Lightweight Open-Source LLMs for Psychological Defense Mechanism Detection*

Siam Karip and Nahid Hossain

*Team Aurum at MedExACT 2026@ACL: Data Augmentation and Clinical Long-former Fine-Tuning for Medical Decision Extraction*

Jyoti Kumari, Vinay Ulli and Anindita Mondal

*NJUST-KMG at MedGenVidQA 2026: Cascade Multi-modal Alignment with Gaussian Soft Priors for Medical Visual Answer Localization*

Jinglong Li and Yang Yang

*LAMAR-2 at MedGenVidQA 2026: Visual Answer Localization in Medical Videos via Multimodal LLM and Context-Augmented Prompting*

Watcharitpol Sermsrisuwan, Nopporn Lekuthai, Seksan Yoadsanit and Titipat Achakulvisut

*Varja-Dominators at MedGenVidQA 2026: Hybrid Video and Document Retrieval using PubMedBERT, T5 Query Expansion, and Cross-Encoder Re-Ranking*

Pratik Dhaktode, Suhani Bighane and Anupama Phakatkar

*Pride-Boiler at MedGenVidQA 2026: LLM-Augmented BM25 Retrieval with Corrective Self-Verification for Biomedical Evidence Retrieval*

Basil Ebinesar, Keyuan Jiang, Charansai Maddineni and Ashok Raja

*Seahawk at MedGenVidQA 2026: LLM Segment-Range Selection for Medical Visual Answer Localization*

Xiaotian Tian and Gulustan Dogan

*UNCC at MedGenVidQA 2026: Structured Temporal Grounding for Medical Video Question Answering*

Hilmi Demirhan and Wlodek Zadrozny

**Saturday, July 4, 2026 (continued)**