

UncertaiNLP 2026

Workshop on Uncertainty-Aware NLP (UncertaiNLP 2026)

Proceedings of the Workshop

October 29, 2026

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



©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 979-8-89176-486-6

Introduction

The way humans use language is inserted in a landscape of nuance and ambiguities that require decoding complex contextual dependencies. For too long, the predominant paradigm in Natural Language Processing (NLP) has worked under the assumption that this richness and underspecifications ought to be collapsed into single, deterministic outputs. The third edition of the Uncertainty-Aware NLP workshop (UncertainNLP 2026), held together with EMNLP in Budapest, challenges this status quo, providing a platform for research that places uncertainty at the centre of linguistic modelling. It embraces variability in human language and aims to represent and evaluate the uncertainty that arises both from language itself and from the modelling tools we use to process it.

Building on the success of previous editions of UncertainNLP at EACL 2024 and EMNLP 2025, this volume reflects the field’s rapid growth and the increased recognition that reliable variability estimation is essential for developing trustworthy NLP. From formal tools and probabilistic inference to the mitigation of hallucinations in large language models, the papers collected here address both epistemic and aleatoric sources of uncertainty across model design, data collection, and evaluation. Together, these contributions help us advance a more nuanced understanding of how uncertainty can be harnessed to improve models, particularly in low-resource settings and mission-critical applications where acknowledging what a system does not know is as important as what it can predict.

This volume contains the proceedings of the third edition of the UncertainNLP workshop hosted on October 29th, 2026, co-located with the 2026 Conference on Empirical Methods in Natural Language Processing at the HUNGEXPO Budapest Congress and Exhibition Center, in Hungary. We invited paper submissions addressing a broad variety of topics, including representing, documenting or modelling uncertainty, parameter estimation, probabilistic inference, decision making, evaluation and calibration, hallucinations and uncertainty-driven mitigation, and uncertainty in multimodal and vision-and-language systems. In addition to the regular workshop submissions, UncertainNLP 2026 joined forces with the *SHROOM initiative to host a shared task. SHROOM-Visions is the fourth edition of the SHROOM shared task on hallucination detection, and is designed to advance model-agnostic evaluation of hallucination detection for vision-and-language models. Building on the SHROOM shared tasks held in 2024 and 2025 and on the recently introduced SHEEP dataset, SHROOM-Visions focused on detecting and classifying fine-grained hallucination spans in image-conditioned text generation in open ended visual question answering. The task used a five-class hallucination taxonomy and covered four languages: Chinese, English, French, and Italian. The shared task attracted substantial participation, with 27 teams contributing more than 600 system submissions.

In total, UncertainNLP 2026 received 243 submissions across three routes: 182 direct submissions, 40 ACL Rolling Review (ARR) commitments, and 21 SHROOM-Visions shared-task papers. The final workshop program comprises 88 contributions: 64 archival papers published in this volume and 24 non-archival presentations. The breakdown by route is as follows.

- The direct route contributed 43 papers to the final program: 32 archival out of 134 archival submissions (23.9%) and 11 non-archival out of 48 non-archival submissions (22.9%). Across both tracks, this represents 23.6% of the 182 submissions received.
- The ARR route contributed 24 papers: 11 archival out of 17 archival commitments (64.7%) and 13 non-archival out of 17 non-archival commitments (76.5%). Six additional commitments had no archival status recorded and did not enter the final program. Across all ARR commitments, the 24 papers represent 60.0% of the 40 commitments received.
- The SHROOM-Visions shared task contributed 21 papers, all archival.

These percentages use all submissions received, including subsequent withdrawals and desk rejections, as their denominators. Across all three routes, the 88 final contributions represent 36.2% of the 243 submissions received.

Across these routes, the proceedings bring together research papers, shared-task system descriptions, and other contributions addressing uncertainty and reliability in NLP and multimodal language technologies.

We are grateful to our invited keynote speakers: Dan Roth (Oracle AI and the University of Pennsylvania), Yingzhen Li (Imperial College London, UK), André Martins (Instituto Superior Técnico, PT) and Gintare Karolina Dziugaite (Google DeepMind, CAN). We would like to acknowledge the support of the Munich Center for Machine Learning (MCML) and of the ERC project Foundation for Empirical Multimodality Research (FOUNDATIONS, agreement No. 101122047).

The UncertaiNLP organizers,

Wilker Aziz, Jonathan Berant, Bryan Eikema, Evgenia Ilia, Marie-Catherine de Marneffe, Barbara Plank, Artem Shelmanov, Swabha Swayamdipta, Jörg Tiedemann, Artem Vazhentsev, Raúl Vázquez, Chrysoula Zerva.

Program Committee

Program Chairs

Wilker Aziz, University of Amsterdam, Netherlands
Jonathan Berant, Tel Aviv University and Google DeepMind, Israel
Bryan Eikema, University of Amsterdam, Netherlands
Evgenia Ilia, University of Amsterdam, Netherlands
Marie-Catherine de Marneffe, UCLouvain and FNRS, Belgium
Barbara Plank, LMU München and IT University of Copenhagen, Germany and Denmark
Artem Shelmanov, Mohamed bin Zayed University of Artificial Intelligence, United Arab Emirates
Swabha Swayamdipta, University of Southern California, Viterbi School of Engineering, United States
Jörg Tiedemann, University of Helsinki, Finland
Artem Vazhentsev, Mohamed bin Zayed University of Artificial Intelligence, United Arab Emirates
Raúl Vázquez, University of Helsinki, Finland
Chrysoula Zerva, Instituto de Telecomunicações, Portugal

Program Committee

Luigi Acerbi, University of Helsinki
Nitay Calderon, Technion
Julius Cheng, University of Cambridge
Ye-eun Cho, Sungkyunkwan University
Caio Corro, INSA Rennes
Vivek Datla, Capital One
Bonaventure Dossou, McGill University
Adam Faulkner, Microsoft
Pedro Ferreira, University of Amsterdam
Jes Frellsen, TU Denmark
Yuu Jinnai, ShizukuAI
Lucie Kunitomo-Jacquín, AIST
Gleb Kuzmin, RUDN University
Yunho Maeng, Ewha Womans University
Yan Meng, University of Amsterdam
Timothee Mickus, University of Helsinki
Tatiana Passali, Aristotle U. of Thessaloniki
Laura Perez-Beltrachini, University of Edinburgh
Yuval Pinter, Ben-Gurion University of the Negev
Timothy Pistotti, University of Auckland
Alberto Purpura, Capital One
Lyudmila Rvanova, Sirius University
Arnab Sharma, University of Paderborn
Anthony Sicilia, West Virginia University
Edwin Simpson, University of Bristol
Sergey Troshin, University of Amsterdam
Grigorios Tsoumakas, Aristotle University of Thessaloniki
Teemu Vahtola, University of Helsinki

Matias Valdenegro, Univ. of Groningen
Sami Virpioja, University of Helsinki
Christian Hardmeier, IT University of Copenhagen
Samuel Barry, Stanford University
Dennis Ulmer, University of Amsterdam
Rico Sennrich, University of Zurich
Oscar Lithgow Serrano, IDSIA
Haau-Sing Li, TU Darmstadt
Nico Daheim, TU Darmstadt
Wataru Hashimoto, NAIST
Alessandro Antonucci, IDSIA
Bidit Das, Amazon
Eshaan Jain, University of Southern California
Zirui Wei, C3.ai
Vladislav Smirnov, MBZUAI
Abhishek Maligehalli Shivalingaiah, Amazon
Pulkit Arya, Pointer
Humayun Akhtar, Prime Robotics
Zihao Zheng, PDD Holdings
Plawan Kumar Rath, Meta
Krishna Aswani, Amazon

Publication Chairs

Raúl Vázquez

Invited Speakers

Dan Roth, Oracle AI and the University of Pennsylvania, USA
Yingzhen Li, Imperial College London, UK
André Martins, Instituto Superior Técnico, Portugal
Gintare Karolina Dziugaite, Google DeepMind, Canada

Keynote Talk

Sparsity Meets Uncertainty: Applications to Conformal Prediction and Generalized Variational Bayes

André Martins

Instituto Superior Técnico, Portugal

29 October 2026 – Time: 09:10–09:50 (Budapest time) – Room: F6

Abstract: In the first part of the talk, I will present a framework for conformal prediction based on temperature scaling of sparse activations. Conformal prediction is a distribution-free framework for uncertainty quantification that replaces point predictions with sets, offering marginal coverage guarantees. We uncover a novel connection between conformal prediction and sparse softmax-like transformations, such as sparsemax and γ -entmax (with $\gamma > 1$), which may assign nonzero probability only to a subset of labels. We introduce new non-conformity scores for classification that make the calibration process correspond to the widely used temperature scaling method. At test time, applying these sparse transformations with the calibrated temperature leads to a support set (i.e., the set of labels with nonzero probability) that automatically inherits the coverage guarantees of conformal prediction.

In the second part, I will present a generalized variational Bayes method based on Fenchel-Young (FY) regularization. Our proposed formulation – FY variational learning – includes as key ingredients new notions of FY free energy, FY evidence, FY evidence lower bound, and FY posterior. We derive alternating minimization and gradient backpropagation algorithms to compute (or lower bound) the FY evidence, which enables learning a wider class of models than previous variational formulations. For a suitable choice of Fenchel-Young regularizer we can obtain sparse and varying support posteriors.

Joint work with Margarida Campos, Sophia Sklaviadis, Mário A.T. Figueiredo, Thomas Moellenhoff, Emtiyaz Khan

Bio: André Martins is an Associate Professor at Instituto Superior Técnico (in both the Computer Science and Electrical and Computer Engineering departments), VP of AI Research at Unbabel, and Senior Researcher at the Instituto de Telecomunicações, where he leads the SARDINE Lab. His research interests span natural language processing and machine learning, with a particular focus on sparse and structured transformations, uncertainty quantification, interpretability, and multimodal processing applied to machine translation, natural language generation, quality estimation, and evaluation. He completed his PhD in the joint CMU–Portugal program in Language Technologies, working at Carnegie Mellon University and Instituto Superior Técnico. His research has been funded by two European Research Council grants — a Starting Grant (DeepSPIN) and a Consolidator Grant (DECOLLAGE). He received multiple best paper awards at ACL conferences, won the WMT 2024 General Translation shared task (achieving the best results in 8 out of 11 languages), and received the IBM Scientific Prize for his work on structured prediction. He is a Fellow of the ELLIS society, co-director of the ELLIS Program in Natural Language Processing, a member of the Lisbon Academy of Sciences, and Program Co-Chair of ACL 2024. He co-founded and co-organizes the Lisbon Machine Learning School (LxMLS).

Keynote Talk

Dan Roth

Oracle AI and the University of Pennsylvania

29 October 2026 – Time: **09:50–10:30 (Budapest time)** – Room: **F6**

Bio: Dan Roth is the Eduardo D. Glandt Distinguished Professor at the University of Pennsylvania and Chief AI Scientist at Oracle. Until June 2024, Dan served as VP and Distinguished Scientist at AWS AI, where he led the scientific effort behind Amazon’s first-generation GenAI products, including Titan Models, Amazon Q, and Amazon Bedrock. His research focuses on the computational foundations of intelligent behavior, spanning machine learning, natural language understanding, knowledge representation and reasoning, and learning theory. He is particularly known for his work on constrained conditional models, integer linear programming formulations for joint learning and inference, and the notion of incidental supervision. Dan is a Fellow of the AAAS, ACM, AAAI, and ACL, and a recipient of the IJCAI John McCarthy Award for major conceptual and theoretical advances in the modeling of natural language understanding, machine learning, and reasoning. He served as Editor-in-Chief of the Journal of Artificial Intelligence Research (JAIR) and has chaired major conferences in his research areas. He received his B.A. Summa cum laude in Mathematics from the Technion, Israel, and his Ph.D. in Computer Science from Harvard University in 1995.

Keynote Talk

Quantifying uncertainty in-context via probing questions

Yingzhen Li

Imperial College London, UK

29 October 2026 – Time: 16:00–16:40 (Budapest time) – Room: F6

Abstract: As large language models (LLMs) gain popularity in conducting prediction tasks in-context, understanding and estimating uncertainty in in-context learning becomes essential to ensuring reliability of prediction and decision-making. The recent hypothesis of in-context learning performing predictive Bayesian inference opens the avenue for Bayesian uncertainty estimation and decision making algorithms. This talk will discuss our recent efforts in decomposing uncertainty and decision making under uncertainty in-context. We first introduce a variational uncertainty decomposition framework for in-context learning without explicitly sampling from the latent parameter posterior, by optimising auxiliary question as probes to obtain an upper bound to the aleatoric uncertainty of an LLM’s in-context learning procedure. We then show the classical Thompson Sampling algorithm can be adapted in this auxiliary question probing framework, where different aggregation strategies for the probes result in different exploration-exploitation trade-off behaviours.

Bio: Yingzhen Li is an Associate Professor in Machine Learning at the Department of Computing, Imperial College London, and a Turing Fellow at the Alan Turing Institute (since March 2024). Her research focuses on building reliable machine learning systems through probabilistic modelling and representation learning, with topics including deep probabilistic graphical model design, fast and accurate Bayesian inference, uncertainty quantification, and robust and adaptive machine learning. Before returning to academia, she spent two and a half years as a Senior Researcher at Microsoft Research Cambridge. She completed her PhD at the University of Cambridge under the supervision of Prof. Richard E. Turner, with a thesis on approximate inference. She has co-organized major events in probabilistic ML, including AABI and AISTATS, and has delivered tutorials at NeurIPS 2020, UAI 2025, and other venues on approximate inference and Bayesian neural networks.

Keynote Talk

Remembering and Forgetting: The Antithetical Roles of Linear Mode Connectivity

Gintare Karolina Dziugaite

Google DeepMind, Canada

29 October 2026 – Time: 16:40–17:20 (Budapest time) – Room: F6

Abstract: Continual learning, model merging, and information unlearning address critical challenges in managing foundation models. In this talk I will present evidence that linear mode connectivity (LMC)—the lack of loss barriers on the linear path connecting two models—is a decisive factor in the success and failure of methods in these three areas. We demonstrate how LMC enables efficient, training-free model merging and continual learning through weight interpolation. Conversely, in unlearning, we show that model editing fails when the edited model is linearly connected to the original model. Attempts to “unlearn” information are often superficial, as the original knowledge remains easily recoverable along the same low-loss path, compromising the model’s robustness and leaving it susceptible to relearning attacks. This work highlights a core trade-off in modern AI: the very properties that make models continually learn also make them difficult to truly and safely edit.

Bio: Gintarė Karolina Dziugaite is a Senior Research Scientist at Google DeepMind, an Adjunct Professor at McGill University, and an Associate Industry Member at Mila. Named a 2019 “Rising Star in Machine Learning,” she has made significant contributions to our understanding of deep learning dynamics. Dziugaite is well known for seminal work on information-theoretic generalization and memorization, and applications to neural networks. She has also made numerous contributions to the science of deep learning, coining the term “linear mode connectivity” and uncovering its deep ties to the Lottery Ticket Hypothesis, optimization landscapes, and sparsity. Her recent work tackles the challenges of data influence and model unlearning, answering critical questions about how models can “forget” or adapt—a central theme for the future of lifelong learning agents.

Table of Contents

<i>Confident in a Confidence Score: Investigating the Sensitivity of the Calibration of Confidence Scores to Supervised Fine-Tuning in Generation Tasks</i>	
Lorenzo Jaime Yu Flores, Cesare Spinoso-Di Piano and Jackie Chi Kit Cheung	1
<i>Tokenization Risk as a Control Signal for Adaptive RAG</i>	
Norbert Braunschweiler, Rama Doddipatla and Nigel Collier	19
<i>AttentionCD: Reducing Factuality Hallucinations in Large Language Models using Task-specific Head-based Contrastive Decoding</i>	
Viju Sudhi, Regina Stodden and Philipp Cimiano	30
<i>Self-Aware Knowledge Probing: Evaluating Language Models' Relational Knowledge through Confidence Calibration</i>	
Christopher Kissling, Elena Merdjanovska and Alan Akbik	48
<i>STEP: Soft Trajectory Estimation of Prediction Correctness via Per-Step Probability Scoring</i>	
Chenyang Yu, Zifan Zeng, Qunli Zhang and Feng Liu	61
<i>Calibration as a First-Class Criterion in LLM Evaluation</i>	
Mario Sanz-Guerrero and Katharina von der Wense	73
<i>Idiomatic Blind Spots in Discrete Diffusion Models</i>	
Joshua Chaz Rivera, Vincent Gripon and Bastien Pasdeloup	81
<i>Trajectory-Triggered Sparse Collaborative Decoding</i>	
Amit Ranjan Trivedi, Sruthikesh Surineni, Kyong Ho Lee, Mohit Mittal, Steven Li and Rex Hill	94
<i>SelectED: Selective LLM Refinement for Entity Disambiguation</i>	
Susanna Rücker, Alexander Reinicke and Alan Akbik	111
<i>Rethinking Aleatoric Uncertainty: Partitioning Irreducible States for Large Language Model Agents</i>	
Renato Okabayashi Miyaji and Matias Valdenegro-Toro	138
<i>Semantic Calibration Prevails Where Token Confidence Fails: Benchmarking Long-Form Scientific QA</i>	
Philip Müller, Nicholas Popovič, Michael Färber and Peter Steinbach	162
<i>The Variance of a Judge: Zero-Variance Rewards by Deterministic Regression</i>	
Kabir Murjani, Parth Vyas, Akshita Abrol, Rajesh Gupta, Sudeep Tanwar and Zhengkui Wang	201
<i>Detection of parametric or contextual knowledge: are the schemas similar across modalities?</i>	
Binesh Arakkal Remesh, Thomas Gerald, Camille Guinaudeau, Sahar Ghannay and Eva Feillet	219
<i>Credal Ensembles for Uncertainty Quantification in LLMs</i>	
Vani Kanjirangat, Oscar William Lithgow-Serrano and Alessandro Antonucci	235
<i>Calibration Is Not Enough: Evaluating Confidence Estimation Under Language Variations</i>	
Yuxi Xia, Dennis Ulmer, Terra Blevins, Yihong Liu, Hinrich Schuetze and Benjamin Roth	245

<i>Explicit State Gating and Implicit Refusal for Unreadable Scene Text</i>	
Dayeon Jang, Yuseok Bae and Kimin Yun	275
<i>The Temporal Tug-of-War: Visualizing and Detecting RAG Conflicts in Diffusion Models via Trajectory Variance</i>	
Sravan Karthick T, Pranav Darshan, Pranav A, Minal Moharir and Ivan P. Yamshchikov ...	280
<i>FAST-Eval and FAST-LoRA: Uncertainty Evaluation Framework and Low-Rank Adaptation for Finding Attributes with Supporting Text</i>	
Samuel Forster and Inmaculada Arnedillo-Sánchez	289
<i>Two Lawyers, Three Opinions: Can Output-Based LLM Uncertainty Find the Cases Where Experts Disagree?</i>	
Laura Braun, Matthias Aßenmacher, Martin Hohenadler and Goeran Kauermann	301
<i>Reliable but Design-Sensitive: Instrument Uncertainty in LLM Annotation</i>	
Thomas Reiter, Christoph Kern, Fedor Miasnikov, Sofiia Nikolenko, Rob Chew, Stephanie Eckman and Frauke Kreuter	312
<i>Verbalized Uncertainty in Coding LLMs: A Contamination-Aware Study of Confidence Calibration</i>	
Hunter Tuerlings and Matias Valdenegro-Toro	328
<i>The Role of Ambiguity in Error Prediction via Uncertainty Quantification</i>	
Ieva Raminta Staliunaite, James Bishop and Andreas Vlachos	356
<i>Uncertainty Meets Conformance: A Process Mining-based Multi-Objective Evaluation for Clinical Reasoning using LLMs</i>	
Seon Kim, Jeong-woo Lee, Tae Hoon Kong and Jongchan Kim	381
<i>DeepLook: Uncertainty-Guided Branching with Lookahead Confidence for Efficient Reasoning</i>	
Tingxin Yang, Zefeng Wang, Mengyue Wang, Jinhe Bi, Xingcheng Zhou, Volker Tresp and Yunpu Ma	390
<i>But Isn't It All Just Semantics? The Importance of Evaluating Uncertainty, Reliability, and Random Effects</i>	
Katarina Pejcinovic, Marianne Casilio, Robert C Gale, Gerasimos Fergadiotis and Steven Bedrick	408
<i>Constrained Decoding Can Reverse Which Inputs Look Uncertain</i>	
Mayur Parvatikar	420
<i>Overview of SHROOM-Visions 2026: A Shared Task on Hallucination Detection in Large Vision-Language Models</i>	
Raúl Vázquez, Aman Sinha, Chuyuan Li, Artem Shelmanov, Artem Vazhentsev, Claudio Savelli, Eduardo Calò, Emilio Raimond, Stella Frank, Hengyu Luo, Flavio Giobergia, Vincent Segonne, Lorenzo Vaiani, Jörg Tiedemann and Timothee Mickus	427
<i>But what is your honest answer? Aiding LLM-judges with honest alternatives using steering vectors</i>	
Leon Eshuijs, Archie Chaudhury, Alan McBeth and Ethan Nguyen	453
<i>TAME: Token Attribution and Masking for Emergent misalignment</i>	
Md Rayhanul Masud and Md Rizwan Parvez	475
<i>Overlap-Based OOD Detection in Closed-Book Question Answering</i>	
Dries Wedda and Matias Valdenegro-Toro	481

<i>Uncertainty-Aware Fact Checking with LLMs</i>	
Yulin Chen, Shubhalaxmi Mukherjee and Pradeep K. Murukannaiah	495
<i>From Pixels to Tokens: Adaptive LLM Decoding and Reliability Analysis with Otsu Thresholding</i>	
Eleni Partalidou, Tatiana Passali, Chrysoula Zerva, Grigorios Tsoumakas and Sophia Ananiadou	518
<i>Are You Sure You're Sure? On the Impact of Instruction Tuning on Confidence and Lexical Diversity</i>	
Irina Proskurina, Mayank Kumar and Oyindolapo Olabisi Komolafe	544
<i>Paraphrase-Induced Retrieval Instability: Systematic Measurement and the Limits of Inference-Time Mitigation</i>	
Aravind Mandiga, Guoming Li and Jin Lu	555
<i>Large Language Models as Automatic Annotators and Annotation Adjudicators for Fine-Grained Opinion Analysis</i>	
Gaurav Negi, M.A. Waskow, John Philip McCrae, Omnia Zayed and Paul Buitelaar	570
<i>BRAVE: Budget-Aware Reflective AutoPrompting for Efficient Evolutionary Prompt Optimization</i>	
Kulin Nikita, Zhuravlev Viktor, Khairullin Artur, Sergey Muravyov, Ilya Makarov and Daniil Sukhorukov	590
<i>Does Quantization Change What Agreement Means? A Paired Audit of Homogeneous LLM Panels</i>	
Ananto Nayan Bala	618
<i>What Counts as Hallucination? A Scoping Review of Tasks, Symptoms, and Convergence Expectations in LLM Research</i>	
Toshiki Kuramoto, Ryohei Kamei and Jun Suzuki	625
<i>Refer, Infer, or Abstain: Investigating Latent Skill Selection of Large Language Models under Contextual Question Answering Uncertainty</i>	
Emil Kalbaliyev and Kairit Sirts	657
<i>Tracing Recurrent and Emergent Hallucination Patterns in Decomposed Text-to-SQL Generation</i>	
Hayden Moore, Abhinav Verma and Anshuman Dutt	672
<i>Beyond Global Calibration: Analyzing Class-Specific Miscalibration and High-Confidence Errors in Text Classification</i>	
Claudio Moisés Valiense de Andrade, Guilherme Fonseca, Washington Cunha, Celso França, Adriana Silvina Pagano, Leonardo Chaves Dutra da Rocha and Marcos André Gonçalves	683
<i>Confidence Drift During Task-Specific Language Model Fine-Tuning</i>	
Soham Tripathy	693
<i>Moral Entropy: Auditing Bias and Uncertainty in Moral Judgment</i>	
Maciej Skorski	704
<i>Team HalluciNauts at SHROOM-Visions: Anchor-Preserving Dual-Resolution Grounding for Fine-Grained Vision Hallucination Detection</i>	
Khang Nhat Hoang Vo	715
<i>Dynamos at SHROOM-Visions: Metric-Aligned Token Tagging and Ensembles for Multilingual Vision-Language Hallucination-Span Detection</i>	
Muhammad Yamin, Sheikh Abdul Wahid and Ali Arshad	722

<i>Sparse-Positive Hallucination Evaluation Entangles Detection and Prevalence</i>	
Srinivas Harish	729
<i>USP at SHROOM-Visions: Why Informative Proxy Visual Signals Do Not Localize Hallucinated Spans</i>	
Rafael M. Fernandes	734
<i>CAML at SHROOM-Visions: Competency-Aware Intermediate-Layer Readouts for Multilingual Hallucination Localization</i>	
Jiwon Jeong, Kimin Yun and Yuseok Bae	743
<i>Irum at SHROOM-Visions 2026: When One Fix Doesn't Fit All — Category-Aware Hallucination Span Detection</i>	
Irum Shehryar	750
<i>Team Bubus at SHROOM-Visions: Probe–Judge Fusion for Hallucination Span Detection</i>	
Berk Bubus	760
<i>Team Eye-Be-Am at SHROOM-Visions</i>	
Eli Schwartz	771
<i>Vroom-Vroom at SHROOM-Visions: A Multi-Judge Committee for Detecting Hallucinated Spans in Vision-Language Outputs</i>	
Toqeer Ehsan, Nico Penttilä, Richard Schmidt, Arash Hajikhani and Victoria Palacin	780
<i>SmurfCat at SHROOM-Visions: Synthetic Augmentation for Multilingual Span-Level Hallucination Detection</i>	
Weerathep Rattanajaratkul, Alexander Panchenko and Elisei Rykov	789
<i>WINNAR749 at SHROOM-Visions: Model-Agnostic Fine-Grained Hallucination Detection in VL-Ms via QLoRA Approach with Soft, Calibrated Targets</i>	
Ranjit Raut	797
<i>L1cache at SHROOM-Visions: MycoGate-VL for Scorer-Aware Multilingual Hallucination Detection</i>	
Aman Behera	806
<i>Team Bit-by-bit at SHROOM-Visions: Adapt, Calibrate, and Average for Character-Level Hallucination Span Detection</i>	
Ioana Branescu and Mihai Dascalu	816
<i>SKstars at SHROOM-Visions: Agreement-Guided Ensembling of Zero-Shot and LoRA-Adapted Vision–Language Models</i>	
ALI ATHAR, Imran Ahsan and Joon-Yong Jung	823
<i>Team CPS Lab at SHROOM-Visions: Parameter-Efficient Vision-Language Hallucination Detection with Threshold Tuning</i>	
Anika Srivastava, Vishal Kumar, HARSHITA SHARMA, Nitisha Aggarwal, Geetika Jain Saxena, Amit Pundir and Sanjeev Singh	828
<i>SpanCalib-VLM: Calibrated Hallucination Span Detection in Vision-Language Models</i>	
Amanuel Gizachew Abebe and Yasmin Moslem	834
<i>Team hccl-jy at SHROOM-Visions: Confidence Signals from an Unmodified VLM Verifier</i>	
Junyoung Jeon and Bongwon Suh	842

<i>Sckwocky at SHROOM-Visions: Character-Level Hallucination-Span Detection in Image-Conditioned VLMs via Dual-Prompted LoRA Decoders</i>	
Kulakov Mikhail, Ruslan Solikhov and Ivan Bondarenko	849
<i>Team champ at SHROOM-Visions: Transferring Fine-Grained Hallucination Detection with Hidden-State Probes and Backbone Adaptation</i>	
Aygalic Jara–Mikolajczak, Thomas Lavergne and Sophie Rosset	856
<i>TÜRK SAT at SHROOM-Visions: Crowd-Emulating Judge Ensembles for Multilingual Hallucination Span Detection in Vision-Language Models</i>	
Alper Karamanlioglu, Gökhan Yurtalan and Berkan Demirel	863
<i>HalluVision at SHROOM-Vision: Hallucination Span Detection with a Classifier over a Vision-Language Backbone</i>	
Mohammed Sohail Pasha Qureshi, Surya Charan and Alapan Kuila	869

Program

Thursday, October 29, 2026

09:00 - 09:10	<i>Opening Remarks</i>
09:10 - 09:50	<i>Keynote Talk 1: André Martins</i>
09:50 - 10:30	<i>Keynote Talk 2: Dan Roth</i>
10:30 - 10:55	<i>Coffee Break</i>
10:55 - 11:25	<i>Poster lightning round 1</i>
11:25 - 12:30	<i>In-Person Poster Session 1</i>
12:30 - 13:40	<i>Lunch Break</i>
13:40 - 13:55	<i>SHROOM</i>
13:55 - 14:25	<i>Poster lightning round 2</i>
14:25 - 15:30	<i>In-Person Poster Session 2</i>
15:30 - 16:00	<i>Coffee Break</i>
16:00 - 16:40	<i>Keynote Talk 3: Yingzhen Li</i>
16:40 - 17:20	<i>Keynote Talk 4: Gintare Karolina Dziugaite</i>
17:20 - 17:30	<i>Closing Remarks</i>