



ICNLSP 2024

**Proceedings of the 7th
International Conference on
Natural Language and Speech
Processing (ICNLSP-2024)**

19–20 October, 2024



©2024 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 979-8-89176-165-0

<https://www.icnlsp.org/>

Introduction

Welcome to the 7th International Conference on Natural Language and Speech Processing (ICNLSP 2024)!

ICNLSP is an excellent opportunity for researchers to discuss trends in the field of Natural Language Processing, and to publish their results in the field.

Many topics were discussed through the interesting works presented during the two days of the conference: speech recognition, machine translation, text summarization, sentiment analysis, large language models, natural language understanding, language resources, and other topics etc.

The program committee accepted 47 good papers (long and short ones) which is around 45% of the received submissions, thanks to the high-quality level of the reviews.

"An Introduction to Large Language Models" is the title of a keynote presented by Prof. Gérard Chollet who was chairing the first session: large language models, followed by 06 oral sessions, namely: information extraction and summarization, human-machine interaction and conversational AI, advances in native language identification and text classification, audio, ASR, and TTS, speech emotion recognition and speaker verification and diarization, and data representation.

Prof. Gérard Chollet has clarified many mysterious points on LLMs and discussed their weaknesses, including hallucinations, the black-box nature of their decision-making, and concerns surrounding security and privacy. He presented symbolic AI and knowledge graphs (KGs) as potential solutions.

We thank all participants for their presentations and discussions during the conference. We would like to thank Gérard Chollet and Hugues Sansen (Institut Polytechnique de Paris) for preparing for the special LLM session.

We appreciate the contribution of Dr. Daniel Braun (University of Twente), Dr. Koichi Takeuchi (Okayama University), Dr. Mohammed Mediani (United Arab Emirates University), Prof. Hend Al-Khalifa (King Saud University), Prof. Yücel Saygın (Sabancı Üniversitesi) and Prof. Nicolas Ballier (Université Paris Cité) for chairing the conference sessions.

Finally, we are grateful to the program committee members for their efforts and commitments.

Mourad Abbas and Abed Alhakim Freihat

Organizers:

General Chair: Dr. Mourad Abbas

Chair: Dr. Abed Alhakim Freihat

Program Chair: Dr. Mourad Abbas

Program Committee:

Ahmed Abdelali, SDAIA, KSA.

Muhammad Abdul-Mageed, University of British Columbia, Canada.

Hend Al-Khalifa, King Saud University, KSA.

Muhammad Al-Qurishi, Elm, KSA.

Mehmet Fatih Amasyalı, Yildiz Technical University, Turkey.

Yuan An, Drexel University, USA.

Christine Balague, Institut Mines-Télécom Business School, France.

Nicolas Ballier, University Paris Cité, France.

Fayssal Bouarourou, University of Strasbourg, France.

Pierrette Bouillon, University of Geneva, Switzerland.

Daniel Braun, University of Twente, Netherlands.

Hadda Cherroun, Amar Telidji University, Algeria.

Gérard Chollet, CNRS, France.

Najim Dehak, Johns Hopkins University, USA.

Abed Alhakim Freihat, University of Trento, Italy.

Mounim El Yacoubi, Telecom SudParis, France.

Ahmed Guessoum, USTHB, Algeria.

Kais Haddar, Faculty of Sciences of Sfax, Tunisia.

Fouzi Harrag, Ferhat Abbas University, Algeria.

Kristiina Jokinen, University of Helsinki, Finland.

Ken Kawamura, Revelata Inc., USA.

Eric Laporte, Gustave Eiffel University, France.

Mohammed Mediani, UAE University, UAE.

Abdelouahab Moussaoui, Ferhat Abbas University, Algeria.

Hermann Ney, RWTH Aachen University, Germany.

Robert Nickel, Bucknell University, USA.

Ahmed Rafea, American University in Cairo, Egypt.

Hugues Sansen, Institut Polytechnique de Paris, France.

Yücel Saygin, Sabanci University, Turkey.

Thomas Schmidt, University of Regensburg, Germany.

Nasredine Semmar, CEA, France.

Peter Sullivan, University of British Columbia, Canada.

Irina Temnikova, Big Data for Smart Society Institute, Bulgaria.

María Inés Torres, University of the Basque Country, Spain.

Jan Trmal, AppTek, USA.

Nikos Tsourakis, University of Geneva, Switzerland.

Guanqun Yang, Stevens Institute of Technology, USA.

Fayçal Ykhlef, CDTA, Algeria.

Wajdi Zaghouani, HBKU, Qatar.

Zbyněk Zajíc, University of West Bohemia, Czechia.

Invited Speakers:

Prof. Gérard Chollet, CNRS, France.

Invited Talk

Prof. Gérard Chollet was granted a PhD in Computer Science and Linguistics from the University of California, Santa Barbara. He taught at Memphis State University and University of Florida before joining CNRS. In 1981, he took in charge the speech research group of Alcatel. In 1983, he joined a CNRS research unit at ENST (Institut Mines-Telecom). In 1992, he participated to the development of IDIAP, a research laboratory of the 'Fondation Dalle Molle' in Martigny, Switzerland. From 1996 to 2012, he was back full time at ENST. He supervised more than forty doctoral theses. CNRS decided in 2012 to grant him an emeritus status within SAMOVAR (Télécom-SudParis). His main research interests are in phonetics, automatic audio-visual speech processing, spoken dialog systems, multimedia, pattern recognition, biometrics, privacy-preserving digital signal processing, speech pathology and speech training aids. His main publications are available from his Google Scholar Citations profile.



An Introduction to Large Language Models

Prof. Gérard Chollet, CNRS, France.

In the tutorial "An Introduction to Large Language Models," key weaknesses of LLMs were discussed, including hallucinations, the black-box nature of their decision-making, indecisiveness, and concerns surrounding security and privacy. To address these challenges, symbolic AI and knowledge graphs (KGs) were introduced as potential solutions. The integration of LLMs with KGs can enhance transparency and reliability in AI outputs, leveraging structured knowledge to mitigate errors and improve decision-making. A significant focus was placed on an embedded solution developed within the e-ViTA project (<https://www.e-vita.coach/>), specifically designed to address privacy issues. This implementation showcases how KGs can provide a framework for safeguarding sensitive information while still allowing LLMs to function effectively. By utilizing KGs, users can access more interpretable and contextually accurate information without compromising data security. The tutorial emphasized the importance of unifying LLMs and KGs to create more robust AI systems that are not only powerful but also trustworthy and secure. Through this approach, the future of AI can balance innovation with ethical considerations, paving the way for more responsible applications in various domains.

Table of Contents

| | |
|---|-----|
| Leveraging Annotator Disagreement for Text Classification | 1 |
| <i>Jin Xu, Mariët Theune and Daniel Braun</i> | |
| Native language Identification for Arabic Language Learners using Pre-trained Language Models | 11 |
| <i>Mohamed Amine Cheragui, Mourad Abbas and Mohammed Mediani</i> | |
| Generative Adversarial Network based Neural Vocoder for Myanmar End-to-End Speech Synthesis | 19 |
| <i>Aye Mya Hlaing and Win Pa Pa</i> | |
| Detecting ChatGPT-Generated Text with GZIP-KNN: A No-Training, Low-Resource Approach | 25 |
| <i>Matthias Berchtold, Sandra Mitrovic, Davide Andreoletti, Daniele Puccinelli and Omran Ayoub</i> | |
| Data Bias According to Bipol: Men are Naturally Right and It is the Role of Women to Follow Their Lead | 34 |
| <i>Irene Pagliai, Goya van Boven, Tosin Adewumi, Lama Alkhaled, Namrata Gurung, Isabella Södergren and Elisa H Barney Smith</i> | |
| Improved Spoken Emotion Recognition With Combined Segment-Based Processing And Triplet Loss | 47 |
| <i>Dejan Porjazovski, Tamas Grosz and Mikko Kurimo</i> | |
| EEG Signal Analysis for Multimodal Simple Concepts Decoding | 55 |
| <i>Sergio Guillén Jiménez, Lorenzo J. Tardón, Ana M Barbancho and Isabel Barbancho</i> | |
| Modeling Score Estimation for Japanese Essays with Generative Pre-trained Transformers | 64 |
| <i>Boago Okgetheng and Koichi Takeuchi</i> | |
| CliqueCorex: A Self-supervised Clique-based Anchored Topic Model | 74 |
| <i>Sami Diaf</i> | |
| Double Decoder: Improving latency for Streaming End-to-end ASR Models | 83 |
| <i>Riqiang Wang, Shreekantha Nadig, Daniil Kulko, Simon Vandieken, Chia-tien Chang, Seyyed Saeed Sarfjoo and Jonas Robertson</i> | |
| Personalised Abusive Language Detection Using LLMs and Retrieval-Augmented Generation | 92 |
| <i>Tsungcheng Yao, Ernest Foo and Sebastian Binnewies</i> | |
| Large-scale Summarization of Chat Transcripts in the Absence of Annotated Summaries | 99 |
| <i>Pratik K. Biswas</i> | |
| Real-Time Speech-Driven Avatar Animation by Predicting Facial landmarks and Deformation Blend-shapes | 109 |
| <i>Juan Camilo Vasquez-Correa, Santiago Moreno-Acevedo, Ander Gonzalez-Docasal, Aritz Lasarguren, Jone Lòpez, Egoitz Rodriguez and Aitor Álvarez</i> | |
| Speech Emotion Recognition for Call Centers using Self-supervised Models: A Complete Pipeline for Industrial Applications | 119 |
| <i>Juan M. Martín-Doñas, Asier López Zorrilla, Mikel deVelasco, Juan Camilo Vasquez-Correa, Aitor Álvarez, Maria Inés Torres, Paz Delgado, Ane Lazpiur, Blanca Romero and Irati Alkorta</i> | |
| Probing Whisper Predictions for French, English and Persian Transcriptions | 129 |
| <i>Nicolas Ballier, Léa Burin, Behnoosh Namdarzadeh, Sara B Ng, Richard Wright and Jean-Baptiste Yunès</i> | |
| Sawaal: A Framework for Automatic Question Generation in Urdu | 139 |

Maria Rahim and Shakeel Ahmed Khoja

| | |
|--|-----|
| Thonburian Whisper: Robust Fine-tuned and Distilled Whisper for Thai | 149 |
| <i>Zaw Htet Aung, Thanachot Thavornmongkol, Atirut Boribalburephan, Vittavas Tangsriworakan, Knot Pipatsrisawat and Titipat Achakulvisut</i> | |
| Dual-Task Learning for AI-Generated Medical Text Detection and Named Entity Recognition | 157 |
| <i>Saja Al-Dabet, Ban Alomar, Sherzod Turaev and Abdelkader Nasreddine Belkacem</i> | |
| Comparative Analysis of Modality Fusion Approaches for Audio-Visual Person Identification and Verification | 168 |
| <i>Aref Farhadipour, Masoumeh Chapariniya, Teodora Vukovic and Volker Dellwo</i> | |
| PoliTun: Tunisian Political Dataset for Detecting Public Opinions and Categories Orientation | 178 |
| <i>Chayma Fourati, Roua Hammami, Chiraz Latiri and Hatem Haddad</i> | |
| Efficient Few-shot Learning for Multi-label Classification of Scientific Documents with Many Classes | 186 |
| <i>Tim Schopf, Alexander Blatzheim, Nektarios Machner and Florian Matthes</i> | |
| A Comparison of Different Tokenization Methods for the Georgian Language | 199 |
| <i>Beso Mikaberidze, Temo Saghinadze, Guram Mikaberidze, Raphael Kalandadze, Konstantine Pkhakadze, Josef van Genabith, Simon Ostermann, Lonneke van der Plas and Philipp Müller</i> | |
| GemST: Continual Learning for End-to-End Speech-to-Text Translation | 209 |
| <i>Pranav Karande, Balaram Sarkar and Chandresh K Maurya</i> | |
| CASCA: Leveraging Role-Based Lexical Cues for Robust Multimodal Speaker Diarization via Large Language Models | 214 |
| <i>William K Nehrbooss</i> | |
| Deep Information Maximisation to Mitigate Information Loss in Text Independent Speaker Verification | 225 |
| <i>Nipun Thejan Fonseka and Nirmal Sankalana</i> | |
| Context-Aware Question Answering in Urdu | 233 |
| <i>Samreen Kazi and Shakeel Ahmed Khoja</i> | |
| Human and Machine: Language Processing in Translation Tasks | 243 |
| <i>Hening Wang, Leixin Zhang and Ondřej Bojar</i> | |
| Asking the Right Questions: Exploiting Hidden Interactions in a Generative Framework for Multilingual, Multitask Classification | 251 |
| <i>Sebastian-Antonio Toma, Camelia Lemnaru, Vlad Andrei Negru and Rodica Potolea</i> | |
| Resolving Gender Biases in LLMs at Inference Time with Novel Dijkstra’s-based K-Explorers Neural Network Traversal (KeNNT) | 267 |
| <i>Hanav Modasiya</i> | |
| Semantically Enriched Text Generation for QA through Dense Paraphrasing | 279 |
| <i>Timothy Obiso, Bingyang Ye, Kyeongmin Rim and James Pustejovsky</i> | |
| Introducing wwm-german-18k - Can LLMs Crack the Million? (Or Win at Least 500 Euros?) | 287 |
| <i>Matthias Aßenmacher, Luis Karrlein, Philipp Schiele and Christian Heumann</i> | |
| Native Language Identification Improves Authorship Attribution | 297 |
| <i>Ahmet Yavuz Uluslu, Gerold Schneider and Can Yildizli</i> | |

| | |
|---|-----|
| Design and Comparison of Arabic Negotiation Bots Using LLMs versus Seq2Seq Models with Reinforcement Learning | 304 |
| <i>Ahmad Hajj, Yasmine A Abu Adla, Samah Albast, Hazem Hajj, Shady Elbassuoni, Wassim El Hajj and Khaled Shaban</i> | |
| Enhancing LLM-based Arabic Negotiation by Fine Tuning on Dialogue Shortcomings | 319 |
| <i>Yasmine A Abu Adla, Hazem Hajj, Shady Elbassuoni, Khaled Shaban and Wassim El Hajj</i> | |
| The Qiyas Benchmark: Measuring ChatGPT Mathematical and Language Understanding in Arabic | 343 |
| <i>Shahad Al-Khalifa and Hend Al-Khalifa</i> | |
| Analyzing Politeness in Arabic Tweets: A Preliminary Study | 352 |
| <i>Hend Al-Khalifa, Nadia Ghezaiel and Maria Bounnit</i> | |
| Feruzaspeech: A 60 Hour Uzbek Read Speech Corpus with Punctuation, Casing, and Context . . | 360 |
| <i>Anna Povey and Katherine Povey</i> | |
| Bulgarian Grammar Error Correction with Data Augmentation and Machine Translation Techniques | 365 |
| <i>Bozhidar Klouchek and Riza Batista-Navarro</i> | |
| On Barriers to Archival Audio Processing | 377 |
| <i>Peter R Sullivan and Muhammad Abdul-Mageed</i> | |
| Conversational Exploratory Search of Scholarly Publications Using Knowledge Graphs | 384 |
| <i>Phillip Schneider and Florian Matthes</i> | |
| A Hybrid Retrieval Approach for Advancing Retrieval-Augmented Generation Systems | 397 |
| <i>Nguyen Nam Doan, Aki Härmä, Remzi Celebi and Valeria Gottardo</i> | |
| Investigating Gender Bias in Large Language Models Through Text Generation | 410 |
| <i>Shweta Soundararajan and Sarah Jane Delany</i> | |
| Improving Long-term F0 representation using post-processing techniques | 425 |
| <i>Crisron Rudolf Lucas, Diptasree Debnath and Andrew Hines</i> | |
| Text-to-Speech in Azerbaijani Language via Transfer Learning in a Low Resource Environment . | 434 |
| <i>Dzhavidan Zeinalov, Bugra Sen and Firangiz Aslanova</i> | |
| SG-RAG: Multi-Hop Question Answering With Large Language Models Through Knowledge Graphs | 439 |
| <i>Ahmmad O. M. Saleh, Gokhan Tur and Yucel Saygin</i> | |
| Linking Quran and Hadith Topics in an Ontology using Word Embeddings and Cellfie Plugin . . . | 449 |
| <i>Ibtisam Khalaf Alshammari, Eric Atwell and Mohammad Ammar Alsalka</i> | |
| Medical Information Extraction with Large Language Models | 456 |
| <i>Raffaello Fornasiere, Nicolò Brunello, Vincenzo Scotti and Mark Carman</i> | |