

ArabicNLP 2023

The First Arabic Natural Language Processing Conference

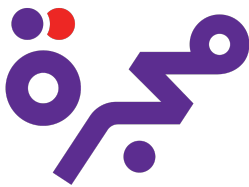
Proceedings of the Conference

December 7, 2023

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



aiXplain



©2023 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 978-1-959429-27-2

Preface

Welcome to The first Arabic Natural Language Processing Conference (ArabicNLP 2023) graduating from the Workshop for Arabic Natural Language Processing Workshop (WANLP) which had its seventh, and last, instance last year, in December 2022 within EMNLP 2022. Over the years, WANLP has developed a growing reputation as a high quality venue for researchers and engineers working on Arabic NLP, where they share and discuss their ongoing work.

The first in the WANLP series was held in Doha, Qatar (EMNLP 2014), followed by Beijing, China (ACL 2015), Valencia, Spain (EACL 2017), Florence, Italy (ACL 2019), online with COLING 2020, online with EACL 2021, then finally a hybrid event in Abu Dhabi, UAE (EMNLP 2022).

For this year’s edition of ArabicNLP, we received a total of 80 main conference submissions and accepted 38 papers (32 long and 6 short), which brings us to an acceptance rate of 47.5%. All papers submitted to the conference were reviewed by at least three reviewers each. Out of the 80 submitted papers, there were 2 desk rejects.

ArabicNLP 2023 included five shared tasks with 48 submissions in totals: (i) The Nuanced Arabic Dialect Identification (NADI) with 13 submissions, (ii) ArAIEval (Persuasion Techniques and Disinformation Detection in Arabic Text) with 17 submissions, (iii) Qur’an QA with 6 submissions, (iv) WojoodNER with 8 submissions, and (v) Arabic Reverse Dictionary with 4 submissions. The shared task overview papers are included in the proceedings. The overview papers and the papers of the shared task winning systems are presented as talks during the conference. None of the shared task papers are counted toward the acceptance rate presented above.

ArabicNLP 2023 also includes a panel discussing the hot topic “Arabic LLMs: Challenges and Opportunities” by leaders in the field, like Areeb Alowisheq, Tom Baldwin, and Kareem Darwish, moderated by Mona Diab.

We were able to secure sponsorship funding from different institutions: King Salman Global Academy for Arabic Language, aiXplain, Lisan.ai, SCAI, Majarra, and Big IR, which we used to support student registrations. We thank all our sponsors for their generous support and their help in building up the Arabic NLP community.

We would like to thank everyone who submitted a paper to the conference, as well as all the members of the Program Committee, who worked hard to provide reviews on a very tight schedule.

Finally, on behalf of everyone involved, organizing committee as well as conference attendees, I would like to thank Nizar Habash and Houda Bouamor for supporting, mentoring and helping this conference be a success, being available for any request and filling any gaps that are overlooked.

Hassan Sawaf, General Chair, on behalf of the conference organizers.

Website of the conference: <https://arabicnlp2023.sigarab.org>

Organizing Committee

General Chair

Hassan Sawaf, aiXplain Inc., USA

Program Chairs

Samhaa El-Beltagy, Newgiza University, Egypt
Wajdi Zaghouani, Hamad Bin Khalifa University, Qatar
Walid Magdy, University of Edinburgh, UK

Publication Chairs

Ahmed Abdelali, National Center for AI, Riyadh, KSA
Nadi Tomeh, Université Sorbonne Paris Nord, France

Shared Tasks Chair

Ibrahim Abu Farha, The University of Sheffield, UK

Scholarships and Awards Chair

Nizar Habash, New York University Abu Dhabi, UAE

Publicity Chairs

Salam Khalifa, Stony Brook University, USA
Amr Keleg, University of Edinburgh, UK

Sponsorship Chairs

Hatem Haddad, University Manouba, Tunisia
Imed Zitouni, Google

Social Chairs

Khalil Mrini, TikTok, USA
Rawan Almatham, ITMSC, King Saud University, KSA

Table of Contents

<i>Violet: A Vision-Language Model for Arabic Image Captioning with Gemini Decoder</i> Abdelrahman Mohamed, Fakhreddin Alwajih, El Moatez Billah Nagoudi, Alcides Alcoba Inciarte and Muhammad Abdul-Mageed	1
<i>Nábra: Syrian Arabic Dialects with Morphological Annotations</i> Amal Nayouf, Tymaa Hasanain Hammouda, Mustafa Jarrar, Fadi A. Zaraket and Mohamad-Bassam Kurdy	12
<i>HICMA: The Handwriting Identification for Calligraphy and Manuscripts in Arabic Dataset</i> Anis Ismail, Zena Kamel and Reem Mahmoud	24
<i>Automated De-Identification of Arabic Medical Records</i> Veysel Kocaman, Youssef MELLAH, Hasham Ul Haq and David Talby	33
<i>ArTST: Arabic Text and Speech Transformer</i> Hawau Olamide Toyin, Amirbek Djanibekov, Ajinkya Kulkarni and Hanan Aldarmaki	41
<i>TARJAMAT: Evaluation of Bard and ChatGPT on Machine Translation of Ten Arabic Varieties</i> Karima Kadaoui, Samar M. Magdy, Abdul Waheed, Md Tawkat Islam Khondaker, Ahmed Oumar El-Shangiti, El Moatez Billah Nagoudi and Muhammad Abdul-Mageed	52
<i>Leveraging Domain Adaptation and Data Augmentation to Improve Qur’anic IR in English and Arabic</i> Vera Pavlova	76
<i>LANS: Large-scale Arabic News Summarization Corpus</i> Abdulaziz Alhamadani, Xuchao Zhang, Jianfeng He, Aadyant Khatri and Chang-Tien Lu	89
<i>Beyond English: Evaluating LLMs for Arabic Grammatical Error Correction</i> Sang Yun Kwon, Gagan Bhatia, El Moatez Billah Nagoudi and Muhammad Abdul-Mageed	101
<i>Aswat: Arabic Audio Dataset for Automatic Speech Recognition Using Speech-Representation Learning</i> Lamy Alkanhal, Abeer Alessa, Elaf Almahmoud and Rana Alaqil	120
<i>Analyzing Multilingual Competency of LLMs in Multi-Turn Instruction Following: A Case Study of Arabic</i> Sabri Boughorbel and Majd Hawasly	128
<i>Cross-Dialectal Named Entity Recognition in Arabic</i> Niama Elkhbir, Urchade Zaratiana, Nadi Tomeh and Thierry Charnois	140
<i>Enhancing Arabic Machine Translation for E-commerce Product Information: Data Quality Challenges and Innovative Selection Approaches</i> Bryan Zhang, Salah Danial and Stephan Walter	150
<i>IDRISI-D: Arabic and English Datasets and Benchmarks for Location Mention Disambiguation over Disaster Microblogs</i> Reem Suwaileh, Tamer Elsayed and Muhammad Imran	158
<i>CamelParser2.0: A State-of-the-Art Dependency Parser for Arabic</i> Ahmed Elshabrawy, Muhammed AbuOdeh, Go Inoue and Nizar Habash	170
<i>GARI: Graph Attention for Relative Isomorphism of Arabic Word Embeddings</i> Muhammad Asif Ali, Maha Alshmrani, Jianbin Qin, Yan Hu and Di Wang	181

<i>ArTrivia: Harvesting Arabic Wikipedia to Build A New Arabic Question Answering Dataset</i> Sultan Alrowili and K Vijay-Shanker	191
<i>ArSarcasMoji Dataset: The Emoji Sentiment Roles in Arabic Ironic Contexts</i> Shatha Ali A. Hakami, Robert Hendley and Phillip Smith	208
<i>Performance Implications of Using Unrepresentative Corpora in Arabic Natural Language Processing</i> Saied Alshahrani, Norah Alshahrani, Soumyabrata Dey and Jeanna Matthews	218
<i>Octopus: A Multitask Model and Toolkit for Arabic Natural Language Generation</i> AbdelRahim Elmadany, El Moatez Billah Nagoudi and Muhammad Abdul-Mageed	232
<i>AlGhafa Evaluation Benchmark for Arabic Language Models</i> Ebtessam Almazrouei, Ruxandra Cojocaru, Michele Baldo, Quentin Malartic, Hamza Alobeidli, Daniele Mazzotta, Guilherme Penedo, Giulia Campesan, Mugariya Farooq, Maitha Alhammadi, Julien Launay and Badreddine Nounne	244
<i>ArBanking77: Intent Detection Neural Model and a New Dataset in Modern and Dialectical Arabic</i> Mustafa Jarrar, Ahmet Birim, Mohammed Khalilia, Mustafa Erden and Sana Ghanem	276
<i>ArabCros: AI-Powered Arabic Crossword Puzzle Generation for Educational Applications</i> Kamyar Zeinalipour, Mohamed Zaky Saad, Marco Maggini and Marco Gori	288
<i>Machine Translation of Omani Arabic Dialect from Social Media</i> Khoulou Al-Kharusi and Abdurahman Khalifa AAlAbdulsalam	302
<i>Arabic Fine-Grained Entity Recognition</i> Haneen Abdallatif Liqreina, Mustafa Jarrar, Mohammed Khalilia, Ahmed Oumar El-Shangiti and Muhammad Abdul-Mageed	310
<i>Investigating Zero-shot Cross-lingual Language Understanding for Arabic</i> Zaid Alyafeai and Moataz Ahmed	324
<i>Evaluating ChatGPT and Bard AI on Arabic Sentiment Analysis</i> Abdulmohsen Al-Thubaity, Sakhar Alkhereyf, Hanan Murayshid, Nouf Alshalawi, Maha Bin Omirah, Raghad Alateeq, Rawabi Almutairi, Razan Alsuwailam, Manal Alhassoun and Imaan Alkhamen	335
<i>In-Context Meta-Learning vs. Semantic Score-Based Similarity: A Comparative Study in Arabic Short Answer Grading</i> Menna Fateen and Tsunenori Mina	350
<i>SALMA: Arabic Sense-Annotated Corpus and WSD Benchmarks</i> Mustafa Jarrar, Sanad Malaysha, Tymaa Hasanain Hammouda and Mohammed Khalilia	359
<i>Arabic dialect identification: An in-depth error analysis on the MADAR parallel corpus</i> Helene Bøsei Olsen, Samia Touileb and Erik Velldal	370
<i>Arabic Dialect Identification under Scrutiny: Limitations of Single-label Classification</i> Amr Keleg and Walid Magdy	385
<i>Arabic Topic Classification in the Generative and AutoML Era</i> Doha Albared, Hadi Hamoud and Fadi A. Zaraket	399
<i>On Enhancing Fine-Tuning for Pre-trained Language Models</i> Abir Betka, Zeyd Ferhat, Riyadh Barka, Selma Boutiba, Zineddine S. Kahhou, Tiar M. Lakhdar, Ahmed Abdelali and Habiba Dahmani	405

<i>Multi-Parallel Corpus of North Levantine Arabic</i> Mateusz Krubiński, Hashem Sellat, Shadi Saleh, Adam Pospíšil, Petr Zemánek and Pavel Pecina 411	
<i>Simplify: Automatic Arabic Sentence Simplification using Word Embeddings</i> Yousef SalahEldin and Caroline Sabty	418
<i>Offensive Language Detection in Arabizi</i> Imene Bensalem, Meryem Ait Mout and Paolo Rosso	423
<i>Yet Another Model for Arabic Dialect Identification</i> Ajinkya Kulkarni and Hanan Aldarmaki	435
<i>VoxArabica: A Robust Dialect-Aware Arabic Speech Recognition System</i> Abdul Waheed, Bashar Talafha, Peter Sullivan, AbdelRahim Elmadany and Muhammad Abdul- Mageed	441
<i>KSAA-RD Shared Task: Arabic Reverse Dictionary</i> Rawan Al-Matham, Waad Alshammari, Abdulrahman AIOsaimy, Sarah Alhumoud, Asma Al Wazrah, Afrah Altamimi, Halah Alharbi and Abdullah Alaifi	450
<i>UWB at Arabic Reverse Dictionary shared task: Computing the meaning of a gloss</i> Stephen Taylor	461
<i>Qamosy at Arabic Reverse Dictionary shared task: Semi Decoder Architecture for Reverse Dictionary with SBERT Encoder</i> Serry Sibae, Samar Ahmad, Ibrahim Khurfan, Vian Sabeeh, Ahmed Bahaaulddin, Hanan M. Belhaj and Abdullah I. Alharbi	467
<i>Abed at KSAA-RD Shared Task: Enhancing Arabic Word Embedding with Modified BERT Multilingual</i> Abdelrahim Qaddoumi	472
<i>Rosetta Stone at KSAA-RD Shared Task: A Hop From Language Modeling To Word–Definition Align- ment</i> Ahmed Elbakry, Mohamed Gabr, Muhammad ElNokrashy and Badr AlKhamissi	477
<i>ArAIEval Shared Task: Persuasion Techniques and Disinformation Detection in Arabic Text</i> Maram Hasanain, Firoj Alam, Hamdy Mubarak, Samir Abdaljalil, Wajdi Zaghoulani, Preslav Na- kov, Giovanni Da San Martino and Abed Alhakim Freihat	483
<i>DetectiveRedasers at ArAIEval Shared Task: Leveraging Transformer Ensembles for Arabic Deception Detection</i> Bryan E. Tuck, Fatima Zahra Qachfar, Dainis Boumber and Rakesh M. Verma	494
<i>HTE at ArAIEval Shared Task: Integrating Content Type Information in Binary Persuasive Technique Detection</i> Khaldi Hadjer and Taqiy Eddine Bouklouha	502
<i>USTHB at ArAIEval’23 Shared Task: Disinformation Detection System based on Linguistic Feature Concatenation</i> Mohamed Lichouri, Khaled Lounnas, Aicha Zitouni, Houda Latrache and Rachida Djeradi . .	508
<i>Mavericks at ArAIEval Shared Task: Towards a Safer Digital Space - Transformer Ensemble Models Tackling Deception and Persuasion</i> Sudeep Mangalvedhekar, Kshitij Deshpande, Yash Patwardhan, Vedant Deshpande and Ravindra Murumkar	513

<i>KnowTellConvince at ArAIEval Shared Task: Disinformation and Persuasion Detection in Arabic using Similar and Contrastive Representation Alignment</i>	
Hariram Veeramani, Surendrabikram Thapa and Usman Naseem	519
<i>PTUK-HULAT at ArAIEval Shared Task Fine-tuned Distilbert to Predict Disinformative Tweets</i>	
Areej Jaber and Paloma Martinez	525
<i>AraDetector at ArAIEval Shared Task: An Ensemble of Arabic-specific pre-trained BERT and GPT-4 for Arabic Disinformation Detection</i>	
Ahmed Bahaaulddin, Vian Sabeeh, Hanan M. Belhaj, Serry Sibae, Samar Ahmad, Ibrahim Khurfan and Abdullah I. Alharbi	530
<i>rematchka at ArAIEval Shared Task: Prefix-Tuning & Prompt-tuning for Improved Detection of Propaganda and Disinformation in Arabic Social Media Content</i>	
Reem Abdel-Salam	536
<i>Itri Amigos at ArAIEval Shared Task: Transformer vs. Compression-Based Models for Persuasion Techniques and Disinformation Detection</i>	
Jehad Oumer, Nouman Ahmed and Natalia Flechas Manrique	543
<i>ReDASPersuasion at ArAIEval Shared Task: Multilingual and Monolingual Models For Arabic Persuasion Detection</i>	
Fatima Zahra Qachfar and Rakesh M. Verma	549
<i>UL & UM6P at ArAIEval Shared Task: Transformer-based model for Persuasion Techniques and Disinformation detection in Arabic</i>	
Salima Lamsiyah, Abdelkader El Mahdaouy, Hamza Alami, Ismail Berrada and Christoph Schommer	558
<i>AAST-NLP at ArAIEval Shared Task: Tackling Persuasion technique and Disinformation Detection using Pre-Trained Language Models On Imbalanced Datasets</i>	
Ahmed El-Sayed, Omar Nasr and Noureldin Elmadany	565
<i>PD-AR at ArAIEval Shared Task: A BERT-Centric Approach to Tackle Arabic Disinformation</i>	
Pritam Deka and Ashwathy T. Revi	570
<i>Nexus at ArAIEval Shared Task: Fine-Tuning Arabic Language Models for Propaganda and Disinformation Detection</i>	
Yunze Xiao and Firoj Alam	576
<i>Frank at ArAIEval Shared Task: Arabic Persuasion and Disinformation: The Power of Pretrained Models</i>	
Dilshod Azizov, Jiyong Li and Shangsong Liang	583
<i>Raphael at ArAIEval Shared Task: Understanding Persuasive Language and Tone, an LLM Approach</i>	
Utsav Shukla, Manan Vyas and Shailendra Tiwari	589
<i>Legend at ArAIEval Shared Task: Persuasion Technique Detection using a Language-Agnostic Text Representation Model</i>	
Olumide E. Ojo, Olaronke O. Adebajji, Hiram Calvo, Damian O. Dieke, Olumuyiwa E. OJO, Seye E. Akinsanya, Tolulope O. Abiola and Anna Feldman	594
<i>NADI 2023: The Fourth Nuanced Arabic Dialect Identification Shared Task</i>	
Muhammad Abdul-Mageed, AbdelRahim Elmadany, Chiyu Zhang, El Moatez Billah Nagoudi, Houda Bouamor and Nizar Habash	600

<i>DialectNLU at NADI 2023 Shared Task: Transformer Based Multitask Approach Jointly Integrating Dialect and Machine Translation Tasks in Arabic</i>	
Hariram Veeramani, Surendrabikram Thapa and Usman Naseem	614
<i>UoT at NADI 2023 shared task: Automatic Arabic Dialect Identification is Made Possible</i>	
Abdulslam F A Nwesri, Nabila A S Shinbir and Hassan A H Ebrahim	620
<i>SANA at NADI 2023 shared task: Ensemble of Layer-Wise BERT-based models for Dialectal Arabic Identification</i>	
Nada Almarwani and Samah Aloufi	625
<i>ISL-AAST at NADI 2023 shared task: Enhancing Arabic Dialect Identification in the Era of Globalization and Technological Progress</i>	
Shorouk Adel and Nouredin Elmadany	631
<i>Frank at NADI 2023 Shared Task: Trio-Based Ensemble Approach for Arabic Dialect Identification</i>	
Dilshod Azizov, Jiyong Li and Shangsong Liang	637
<i>NLPeople at NADI 2023 Shared Task: Arabic Dialect Identification with Augmented Context and Multi-Stage Tuning</i>	
Mohab Elkaref, Movina Moses, Shinnosuke Tanaka, James Barry and Geeth De Mel	642
<i>USTHB at NADI 2023 shared task: Exploring Preprocessing and Feature Engineering Strategies for Arabic Dialect Identification</i>	
Mohamed Lichouri, Khaled Lounnas, Aicha Zitouni, Houda Latrache and Rachida Djeradi . .	647
<i>rematchka at NADI 2023 shared task: Parameter Efficient tuning for Dialect Identification and Dialect Machine Translation</i>	
Reem Abdel-Salam	652
<i>UniManc at NADI 2023 Shared Task: A Comparison of Various T5-based Models for Translating Arabic Dialectal Text to Modern Standard Arabic</i>	
Abdullah Khered, Ingy Yasser Abdelhalim, Nadine Abdelhalim, Ahmed Soliman and Riza Batista-Navarro	658
<i>IUNADI at NADI 2023 shared task: Country-level Arabic Dialect Classification in Tweets for the Shared Task NADI 2023</i>	
Yash A. Hatekar and Muhammad S. Abdo	665
<i>The Helsinki-NLP Submissions at NADI 2023 Shared Task: Walking the Baseline</i>	
Yves Scherrer, Aleksandra Miletić and Olli Kuparinen	670
<i>Mavericks at NADI 2023 Shared Task: Unravelling Regional Nuances through Dialect Identification using Transformer-based Approach</i>	
Vedant Deshpande, Yash Patwardhan, Kshitij Deshpande, Sudeep Mangalvedhekar and Ravindra Murumkar	678
<i>ANLP-RG at NADI 2023 shared task: Machine Translation of Arabic Dialects: A Comparative Study of Transformer Models</i>	
Wiem Derouich, Sameh Kchaou and Rahma Boujelbane	683
<i>Qur'an QA 2023 Shared Task: Overview of Passage Retrieval and Reading Comprehension Tasks over the Holy Qur'an</i>	
Rana Malhas, Watheq Mansour and Tamer Elsayed	690

<i>AHJL at Qur'an QA 2023 Shared Task: Enhancing Passage Retrieval using Sentence Transformer and Translation</i>	
Hessa A. Alawwad, Lujain A. Alawwad, Jamilah Alharbi and Abdullah I. Alharbi	702
<i>LowResContextQA at Qur'an QA 2023 Shared Task: Temporal and Sequential Representation Augmented Question Answering Span Detection in Arabic</i>	
Hariram Veeramani, Surendrabikram Thapa and Usman Naseem	708
<i>GYM at Qur'an QA 2023 Shared Task: Multi-Task Transfer Learning for Quranic Passage Retrieval and Question Answering with Large Language Models</i>	
Ghazaleh Mahmoudi, Yeganeh Morshedzadeh and Sauleh Eetemadi	714
<i>LKAU23 at Qur'an QA 2023: Using Transformer Models for Retrieving Passages and Finding Answers to Questions from the Qur'an</i>	
Sarah Alnefaie, Abdullah N. Alsaleh, Eric Atwell, Mohammad Alsalka and Abdulrahman Altahhan	720
<i>TCE at Qur'an QA 2023 Shared Task: Low Resource Enhanced Transformer-based Ensemble Approach for Qur'anic QA</i>	
Mohammed Alaa Elkomy and Amany Sarhan	728
<i>Al-Jawaab at Qur'an QA 2023 Shared Task: Exploring Embeddings and GPT Models for Passage Retrieval and Reading Comprehension</i>	
Abdulrezzak Zekiye and Fadi Amroush	743
<i>WojoodNER 2023: The First Arabic Named Entity Recognition Shared Task</i>	
Mustafa Jarrar, Muhammad Abdul-Mageed, Mohammed Khalilia, Bashar Talafha, AbdelRahim Elmadany, Nagham Hamad and Alaa' Omar	748
<i>ELYADATA at WojoodNER Shared Task: Data and Model-centric Approaches for Arabic Flat and Nested NER</i>	
Imen Laouirine, Haroun Elleuch and Fethi Bougares	759
<i>Lotus at WojoodNER Shared Task: Multilingual Transformers: Unveiling Flat and Nested Entity Recognition</i>	
Jiyong Li, Dilshod Azizov, Hilal AlQuabeh and Shangsong Liang	765
<i>AlexU-AIC at WojoodNER shared task: Sequence Labeling vs MRC and SWA for Arabic Named Entity Recognition</i>	
Shereen Elkordi, Noha Adly and Marwan Torki	771
<i>UM6P & UL at WojoodNER shared task: Improving Multi-Task Learning for Flat and Nested Arabic Named Entity Recognition</i>	
Abdelkader El Mahdaouy, Salima Lamsiyah, Hamza Alami, Christoph Schommer and Ismail Berrada	777
<i>AlphaBrains at WojoodNER shared task: Arabic Named Entity Recognition by Using Character-based Context-Sensitive Word Representations</i>	
Toqeer Ehsan, Amjad Ali and Ala Al-Fuqaha	783
<i>LIPN at WojoodNER shared task: A Span-Based Approach for Flat and Nested Arabic Named Entity Recognition</i>	
Niama Elkhbir, Urchade Zaratiana, Nadi Tomeh and Thierry Charnois	789
<i>Alex-U 2023 NLP at WojoodNER shared task: AraBINDER (Bi-encoder for Arabic Named Entity Recognition)</i>	
Mariam Hussein, Sarah Khaled, Marwan Torki and Nagwa El-Makky	797

El-Kawaref at WjoodNER shared task: StagedNER for Arabic Named Entity Recognition
Nehal Elkaref and Mohab Elkaref.....803

Program

Thursday, December 7, 2023

09:00 - 09:20 *Welcome Session*

09:20 - 10:30 *Arabic Downstream NLP Tasks*

In-Context Meta-Learning vs. Semantic Score-Based Similarity: A Comparative Study in Arabic Short Answer Grading

Menna Fateen and Tsunenori Mina

Arabic Dialect Identification under Scrutiny: Limitations of Single-label Classification

Amr Keleg and Walid Magdy

Nâbra: Syrian Arabic Dialects with Morphological Annotations

Amal Nayouf, Tymaa Hasanain Hammouda, Mustafa Jarrar, Fadi A. Zaraket and Mohamad-Bassam Kurdy

Arabic Fine-Grained Entity Recognition

Haneen Abdallatif Liqreina, Mustafa Jarrar, Mohammed Khalilia, Ahmed Oumar El-Shangiti and Muhammad Abdul-Mageed

Cross-Dialectal Named Entity Recognition in Arabic

Niama Elkhbir, Urchade Zaratiana, Nadi Tomeh and Thierry Charnois

Leveraging Domain Adaptation and Data Augmentation to Improve Qur'anic IR in English and Arabic

Vera Pavlova

ArabIcros: AI-Powered Arabic Crossword Puzzle Generation for Educational Applications

Kamyar Zeinalipour, Mohamed Zaky Saad, Marco Maggini and Marco Gori

10:30 - 11:00 *Coffee Break*

11:00 - 12:30 *LLMs and Applications*

Evaluating ChatGPT and Bard AI on Arabic Sentiment Analysis

Abdulmohsen Al-Thubaity, Sakhar Alkhereyf, Hanan Murayshid, Nouf Alshawi, Maha Bin Omirah, Raghad Alateeq, Rawabi Almutairi, Razan Alsuwailem, Manal Alhassoun and Imaan Alkhanen

Thursday, December 7, 2023 (continued)

TARJAMAT: Evaluation of Bard and ChatGPT on Machine Translation of Ten Arabic Varieties

Karima Kadaoui, Samar M. Magdy, Abdul Waheed, Md Tawkat Islam Khondaker, Ahmed Oumar El-Shangiti, El Moatez Billah Nagoudi and Muhammad Abdul-Mageed

Analyzing Multilingual Competency of LLMs in Multi-Turn Instruction Following: A Case Study of Arabic

Sabri Boughorbel and Majd Hawasly

Beyond English: Evaluating LLMs for Arabic Grammatical Error Correction

Sang Yun Kwon, Gagan Bhatia, El Moatez Billah Nagoudi and Muhammad Abdul-Mageed

Octopus: A Multitask Model and Toolkit for Arabic Natural Language Generation

AbdelRahim Elmadany, El Moatez Billah Nagoudi and Muhammad Abdul-Mageed

AlGhafa Evaluation Benchmark for Arabic Language Models

Ebtesam Almazrouei, Ruxandra Cojocaru, Michele Baldo, Quentin Malartic, Hamza Alobeidli, Daniele Mazzotta, Guilherme Penedo, Giulia Campesan, Mugarriya Farooq, Maitha Alhammedi, Julien Launay and Badreddine Noun

GARI: Graph Attention for Relative Isomorphism of Arabic Word Embeddings

Muhammad Asif Ali, Maha Alshmrani, Jianbin Qin, Yan Hu and Di Wang

Violet: A Vision-Language Model for Arabic Image Captioning with Gemini Decoder

Abdelrahman Mohamed, Fakhraddin Alwajih, El Moatez Billah Nagoudi, Alcides Alcoba Inciarte and Muhammad Abdul-Mageed

ArTST: Arabic Text and Speech Transformer

Hawau Olamide Toyin, Amirbek Djanibekov, Ajinkya Kulkarni and Hanan Aldarmaki

12:30 - 14:00 *Lunch Break*

14:00 - 14:20 *Arabic Core NLP*

SALMA: Arabic Sense-Annotated Corpus and WSD Benchmarks

Mustafa Jarrar, Sanad Malaysha, Tymaa Hasanain Hammouda and Mohammed Khalilia

CamelParser2.0: A State-of-the-Art Dependency Parser for Arabic

Ahmed Elshabrawy, Muhammed AbuOdeh, Go Inoue and Nizar Habash

Thursday, December 7, 2023 (continued)

14:20 - 14:45 *Shared Tasks Overview*

ArAIEval Shared Task: Persuasion Techniques and Disinformation Detection in Arabic Text

Maram Hasanain, Firoj Alam, Hamdy Mubarak, Samir Abdaljalil, Wajdi Zaghouani, Preslav Nakov, Giovanni Da San Martino and Abed Alhakim Freihat

NADI 2023: The Fourth Nuanced Arabic Dialect Identification Shared Task

Muhammad Abdul-Mageed, AbdelRahim Elmadany, Chiyu Zhang, El Moatez Billah Nagoudi, Houda Bouamor and Nizar Habash

WojoodNER 2023: The First Arabic Named Entity Recognition Shared Task

Mustafa Jarrar, Muhammad Abdul-Mageed, Mohammed Khalilia, Bashar Talafha, AbdelRahim Elmadany, Nagham Hamad and Alaa' Omar

KSAA-RD Shared Task: Arabic Reverse Dictionary

Rawan Al-Matham, Waad Alshammari, Abdulrahman AIOsaimy, Sarah Alhuzmoud, Asma Al Wazrah, Afrah Altamimi, Halah Alharbi and Abdullah Alaifi

Qur'an QA 2023 Shared Task: Overview of Passage Retrieval and Reading Comprehension Tasks over the Holy Qur'an

Rana Malhas, Watheq Mansour and Tamer Elsayed

14:45 - 15:30 *Panel on Arabic LLMs: Challenges and Opportunities (Areeb Alowisheq, Kareem Darwish and Preslav Nakov, moderated by Mona Diab)*

15:30 - 16:00 *Coffee Break*

16:00 - 17:30 *Main Conference Posters (in-Person)*

Enhancing Arabic Machine Translation for E-commerce Product Information: Data Quality Challenges and Innovative Selection Approaches

Bryan Zhang, Salah Danial and Stephan Walter

Multi-Parallel Corpus of North Levantine Arabic

Mateusz Krubiński, Hashem Sellat, Shadi Saleh, Adam Pospíšil, Petr Zemánek and Pavel Pecina

VoxArabica: A Robust Dialect-Aware Arabic Speech Recognition System

Abdul Waheed, Bashar Talafha, Peter Sullivan, AbdelRahim Elmadany and Muhammad Abdul-Mageed

Thursday, December 7, 2023 (continued)

Arabic Topic Classification in the Generative and AutoML Era

Doha Albared, Hadi Hamoud and Fadi A. Zaraket

Yet Another Model for Arabic Dialect Identification

Ajinkya Kulkarni and Hanan Aldarmaki

16:00 - 17:30 *EMNLP Findings Posters (in-Person)*

Filtered Semi-Markov CRF

Urchade Zaratiana, Nadi Tomeh, Niama El Khbir, Pierre Holat and Thierry Charnois

Automatic Pronunciation Assessment - A Review

Yassine El Kheir, Ahmed Ali and Shammur Absar Chowdhury

Data Augmentation Techniques for Machine Translation of Code-Switched Texts: A Comparative Study

Injy Hamed, Nizar Habash and Thang Vu

16:00 - 17:30 *Shared Task Posters (in-Person)*

LIPN at WjoodNER shared task: A Span-Based Approach for Flat and Nested Arabic Named Entity Recognition

Niama Elkhbir, Urchade Zaratiana, Nadi Tomeh and Thierry Charnois

DetectiveRedasers at ArAIEval Shared Task: Leveraging Transformer Ensembles for Arabic Deception Detection

Bryan E. Tuck, Fatima Zahra Qachfar, Dainis Boumber and Rakesh M. Verma

Nexus at ArAIEval Shared Task: Fine-Tuning Arabic Language Models for Propaganda and Disinformation Detection

Yunze Xiao and Firoj Alam

PTUK-HULAT at ArAIEval Shared Task Fine-tuned Distilbert to Predict Disinformative Tweets

Areej Jaber and Paloma Martinez

ReDASPersuasion at ArAIEval Shared Task: Multilingual and Monolingual Models For Arabic Persuasion Detection

Fatima Zahra Qachfar and Rakesh M. Verma

Thursday, December 7, 2023 (continued)

UL & UM6P at ArAIEval Shared Task: Transformer-based model for Persuasion Techniques and Disinformation detection in Arabic

Salima Lamsiyah, Abdelkader El Mahdaouy, Hamza Alami, Ismail Berrada and Christoph Schommer

UWB at Arabic Reverse Dictionary shared task: Computing the meaning of a gloss

Stephen Taylor

NLPeople at NADI 2023 Shared Task: Arabic Dialect Identification with Augmented Context and Multi-Stage Tuning

Mohab Elkaref, Movina Moses, Shinnosuke Tanaka, James Barry and Geeth De Mel

SANA at NADI 2023 shared task: Ensemble of Layer-Wise BERT-based models for Dialectal Arabic Identification

Nada Almarwani and Samah Aloufi

The Helsinki-NLP Submissions at NADI 2023 Shared Task: Walking the Baseline

Yves Scherrer, Aleksandra Miletic and Olli Kuparinen

LKAU23 at Qur'an QA 2023: Using Transformer Models for Retrieving Passages and Finding Answers to Questions from the Qur'an

Sarah Alnefaie, Abdullah N. Alsaleh, Eric Atwell, Mohammad Alsalka and Abdulrahman Altahhan

El-Kawaref at WojooodNER shared task: StagedNER for Arabic Named Entity Recognition

Nehal Elkaref and Mohab Elkaref

UM6P & UL at WojooodNER shared task: Improving Multi-Task Learning for Flat and Nested Arabic Named Entity Recognition

Abdelkader El Mahdaouy, Salima Lamsiyah, Hamza Alami, Christoph Schommer and Ismail Berrada

16:00 - 17:30 *Main Conference Posters (Virtual)*

IDRISI-D: Arabic and English Datasets and Benchmarks for Location Mention Disambiguation over Disaster Microblogs

Reem Suwaileh, Tamer Elsayed and Muhammad Imran

HICMA: The Handwriting Identification for Calligraphy and Manuscripts in Arabic Dataset

Anis Ismail, Zena Kamel and Reem Mahmoud

Thursday, December 7, 2023 (continued)

Simplify: Automatic Arabic Sentence Simplification using Word Embeddings
Yousef SalahEldin and Caroline Sabty

ArBanking77: Intent Detection Neural Model and a New Dataset in Modern and Dialectical Arabic

Mustafa Jarrar, Ahmet Birim, Mohammed Khalilia, Mustafa Erden and Sana Ghanem

ArSarcasMoji Dataset: The Emoji Sentiment Roles in Arabic Ironic Contexts
Shatha Ali A. Hakami, Robert Hendley and Phillip Smith

Machine Translation of Omani Arabic Dialect from Social Media
Khoula Al-Kharusi and Abdurahman Khalifa AAlAbdulsalam

Automated De-Identification of Arabic Medical Records
Veysel Kocaman, Youssef MELLAH, Hasham Ul Haq and David Talby

Arabic dialect identification: An in-depth error analysis on the MADAR parallel corpus
Helene Bøsei Olsen, Samia Touileb and Erik Velldal

Investigating Zero-shot Cross-lingual Language Understanding for Arabic
Zaid Alyafeai and Moataz Ahmed

On Enhancing Fine-Tuning for Pre-trained Language Models
Abir Betka, Zeyd Ferhat, Riyadh Barka, Selma Boutiba, Zineddine S. Kahhoul, Tiar M. Lakhdar, Ahmed Abdelali and Habiba Dahmani

LANS: Large-scale Arabic News Summarization Corpus
Abdulaziz Alhamadani, Xuchao Zhang, Jianfeng He, Aadyant Khatri and Chang-Tien Lu

Performance Implications of Using Unrepresentative Corpora in Arabic Natural Language Processing
Saied Alshahrani, Norah Alshahrani, Soumyabrata Dey and Jeanna Matthews

ArTrivia: Harvesting Arabic Wikipedia to Build A New Arabic Question Answering Dataset
Sultan Alrowili and K Vijay-Shanker

Thursday, December 7, 2023 (continued)

Aswat: Arabic Audio Dataset for Automatic Speech Recognition Using Speech-Representation Learning

Lama Alkanhal, Abeer Alessa, Elaf Almahmoud and Rana Alaql

Offensive Language Detection in Arabizi

Imene Bensalem, Meryem Ait Mout and Paolo Rosso

16:00 - 17:30 *EMNLP Findings Posters (Virtual)*

Arabic Mini-ClimateGPT : A Climate Change and Sustainability Tailored Arabic LLM

Sahal Shaji Mullappilly, Abdelrahman M Shaker, Omkar Chakradhar Thawakar, Hisham Cholakkal, Rao Muhammad Anwer, Salman Khan and Fahad Khan

Dolphin: A Challenging and Diverse Benchmark for Arabic NLG

El Moatez Billah Nagoudi, AbdelRahim A. Elmadany, Ahmed Oumar El-Shangiti and Muhammad Abdul-Mageed

16:00 - 17:30 *Shared Task Posters (Virtual)*

AAST-NLP at AraIEval Shared Task: Tackling Persuasion technique and Disinformation Detection using Pre-Trained Language Models On Imbalanced Datasets

Ahmed El-Sayed, Omar Nasr and Noureldin Elmadany

AraDetector at AraIEval Shared Task: An Ensemble of Arabic-specific pre-trained BERT and GPT-4 for Arabic Disinformation Detection

Ahmed Bahaaulddin, Vian Sabeeh, Hanan M. Belhaj, Serry Sibae, Samar Ahmad, Ibrahim Khurfan and Abdullah I. Alharbi

Frank at AraIEval Shared Task: Arabic Persuasion and Disinformation: The Power of Pretrained Models

Dilshod Azizov, Jiyong Li and Shangsong Liang

HTE at AraIEval Shared Task: Integrating Content Type Information in Binary Persuasive Technique Detection

Khalid Hadjer and Taqiy Eddine Bouklouha

Itri Amigos at AraIEval Shared Task: Transformer vs. Compression-Based Models for Persuasion Techniques and Disinformation Detection

Jehad Oumer, Nouman Ahmed and Natalia Flechas Manrique

KnowTellConvince at AraIEval Shared Task: Disinformation and Persuasion Detection in Arabic using Similar and Contrastive Representation Alignment

Hariram Veeramani, Surendrabikram Thapa and Usman Naseem

Thursday, December 7, 2023 (continued)

Legend at ArAIEval Shared Task: Persuasion Technique Detection using a Language-Agnostic Text Representation Model

Olumide E. Ojo, Olaronke O. Adebajji, Hiram Calvo, Damian O. Dieke, Olumuyiwa E. OJO, Seye E. Akinsanya, Tolulope O. Abiola and Anna Feldman

Mavericks at ArAIEval Shared Task: Towards a Safer Digital Space - Transformer Ensemble Models Tackling Deception and Persuasion

Sudeep Mangalvedhekar, Kshitij Deshpande, Yash Patwardhan, Vedant Deshpande and Ravindra Murumkar

PD-AR at ArAIEval Shared Task: A BERT-Centric Approach to Tackle Arabic Disinformation

Pritam Deka and Ashwathy T. Revi

Raphael at ArAIEval Shared Task: Understanding Persuasive Language and Tone, an LLM Approach

Utsav Shukla, Manan Vyas and Shailendra Tiwari

USTHB at NADI 2023 shared task: Exploring Preprocessing and Feature Engineering Strategies for Arabic Dialect Identification

Mohamed Lichouri, Khaled Lounnas, Aicha Zitouni, Houda Latrache and Rachida Djeradi

rematchka at ArAIEval Shared Task: Prefix-Tuning & Prompt-tuning for Improved Detection of Propaganda and Disinformation in Arabic Social Media Content

Reem Abdel-Salam

Abed at KSAA-RD Shared Task: Enhancing Arabic Word Embedding with Modified BERT Multilingual

Abdelrahim Qaddoumi

Qamosy at Arabic Reverse Dictionary shared task: Semi Decoder Architecture for Reverse Dictionary with SBERT Encoder

Serry Sibae, Samar Ahmad, Ibrahim Khurfan, Vian Sabeeh, Ahmed Bahaauldin, Hanan M. Belhaj and Abdullah I. Alharbi

Rosetta Stone at KSAA-RD Shared Task: A Hop From Language Modeling To Word-Definition Alignment

Ahmed Elbakry, Mohamed Gabr, Muhammad ElNokrashy and Badr AlKhamissi

ANLP-RG at NADI 2023 shared task: Machine Translation of Arabic Dialects: A Comparative Study of Transformer Models

Wiem Derouich, Sameh Kchaou and Rahma Boujelbane

DialectNLU at NADI 2023 Shared Task: Transformer Based Multitask Approach Jointly Integrating Dialect and Machine Translation Tasks in Arabic

Hariram Veeramani, Surendrabikram Thapa and Usman Naseem

Thursday, December 7, 2023 (continued)

Frank at NADI 2023 Shared Task: Trio-Based Ensemble Approach for Arabic Dialect Identification

Dilshod Azizov, Jiyong Li and Shangsong Liang

ISL-AAST at NADI 2023 shared task: Enhancing Arabic Dialect Identification in the Era of Globalization and Technological Progress

Shorouk Adel and Noureldin Elmadany

IUNADI at NADI 2023 shared task: Country-level Arabic Dialect Classification in Tweets for the Shared Task NADI 2023

Yash A. Hatekar and Muhammad S. Abdo

Mavericks at NADI 2023 Shared Task: Unravelling Regional Nuances through Dialect Identification using Transformer-based Approach

Vedant Deshpande, Yash Patwardhan, Kshitij Deshpande, Sudeep Mangalvedhekar and Ravindra Murumkar

USTHB at NADI 2023 shared task: Exploring Preprocessing and Feature Engineering Strategies for Arabic Dialect Identification

Mohamed Lichouri, Khaled Lounnas, Aicha Zitouni, Houda Latrache and Rachida Djeradi

UniManc at NADI 2023 Shared Task: A Comparison of Various T5-based Models for Translating Arabic Dialectical Text to Modern Standard Arabic

Abdullah Khered, Ingy Yasser Abdelhalim, Nadine Abdelhalim, Ahmed Soliman and Riza Batista-Navarro

UoT at NADI 2023 shared task: Automatic Arabic Dialect Identification is Made Possible

Abduslam F A Nwesri, Nabila A S Shinbir and Hassan A H Ebrahim

rematchka at NADI 2023 shared task: Parameter Efficient tuning for Dialect Identification and Dialect Machine Translation

Reem Abdel-Salam

AHJL at Qur'an QA 2023 Shared Task: Enhancing Passage Retrieval using Sentence Transformer and Translation

Hessa A. Alawwad, Lujain A. Alawwad, Jamilah Alharbi and Abdullah I. Alharbi

Al-Jawaab at Qur'an QA 2023 Shared Task: Exploring Embeddings and GPT Models for Passage Retrieval and Reading Comprehension

Abdulrezzak Zekiye and Fadi Amroush

GYM at Qur'an QA 2023 Shared Task: Multi-Task Transfer Learning for Quranic Passage Retrieval and Question Answering with Large Language Models

Ghazaleh Mahmoudi, Yeganeh Morshedzadeh and Sauleh Eetemadi

Thursday, December 7, 2023 (continued)

LowResContextQA at Qur'an QA 2023 Shared Task: Temporal and Sequential Representation Augmented Question Answering Span Detection in Arabic

Hariram Veeramani, Surendrabikram Thapa and Usman Naseem

Alex-U 2023 NLP at WojooodNER shared task: AraBINDER (Bi-encoder for Arabic Named Entity Recognition)

Mariam Hussein, Sarah Khaled, Marwan Toriki and Nagwa El-Makky

AlexU-AIC at WojooodNER shared task: Sequence Labeling vs MRC and SWA for Arabic Named Entity Recognition

Shereen Elkordi, Noha Adly and Marwan Toriki

AlphaBrains at WojooodNER shared task: Arabic Named Entity Recognition by Using Character-based Context-Sensitive Word Representations

Toqeer Ehsan, Amjad Ali and Ala Al-Fuqaha

ELYADATA at WojooodNER Shared Task: Data and Model-centric Approaches for Arabic Flat and Nested NER

Imen Laouirine, Haroun Elleuch and Fethi Bougares

Lotus at WojooodNER Shared Task: Multilingual Transformers: Unveiling Flat and Nested Entity Recognition

Jiyong Li, Dilshod Azizov, Hilal AlQuabeh and Shangsong Liang

TCE at Qur'an QA 2023 Shared Task: Low Resource Enhanced Transformer-based Ensemble Approach for Qur'anic QA

Mohammed Alaa Elkomy and Amany Sarhan