



**Arabic Natural Language Processing Workshop
co-located with EMNLP 2014, Doha, Qatar**

Second Call for Papers and Participation

Arabic Natural Language Processing Workshop co-located with EMNLP 2014, Doha, Qatar

Workshop date: Saturday October 25, 2014

Extended Paper submission deadline: July 28, 2014

Shared task registration deadline: July 1, 2014

****NEW** The workshop talks are now on Youtube!**

Workshop proceedings are online here

Submission instructions are available here

List of Accepted Papers

Workshop Schedule

WORKSHOP DESCRIPTION

There has been a lot of progress in the last 15 years in the area of Arabic Natural Language Processing (NLP). Many Arabic NLP (or Arabic NLP-related) workshops and conferences have taken place, both in the Arab World and in association with international conferences. This workshop follows in the footsteps of previous efforts to provide a forum for researchers to share and discuss their ongoing work.

We invite submissions on topics that include, but are not limited to, the following:

- Basic core technologies: morphological analysis, disambiguation, tokenization, POS tagging, named entity detection, chunking, parsing, semantic role labeling, sentiment analysis, Arabic dialect modeling, etc.
- Applications: machine translation, speech recognition, speech synthesis, optical character recognition, pedagogy, assistive technologies, social media, etc.
- Resources: dictionaries, annotated data, specialized databases etc.

Submissions may include work in progress as well as finished work. Submissions must have a clear focus on specific issues pertaining to the Arabic language whether it is standard Arabic, dialectal, or mixed. Descriptions of commercial systems are welcome, but authors should be willing to discuss the details of their work. Submissions are expected to be 8 pages long plus 2 pages for references.

IMPORTANT DATES

Paper submission deadline: July 28, 2014

Author notification: August 26, 2014

Camera Ready: September 15, 2014

Workshop: October 25, 2014

SHARED TASK

Associated with the workshop, there will also be a shared task on Arabic text error correction, which we provide further details about [here](#).

ORGANIZERS

Program Co-chairs

Nizar Habash, New York University Abu Dhabi
Stephan Vogel, Qatar Computing Research Institute

Publication Co-chairs

Nadi Tomeh, Paris 13 University
Houda Bouamor, Carnegie Mellon University Qatar

Website Committee

Noura Farra, Columbia University
Kareem Darwish, Qatar Computing Research Institute

Shared Task Committee

Behrang Mohit (co-chair), Carnegie Mellon University Qatar
Alla Rozovskaya (co-chair), Columbia University
Wajdi Zaghouni, Carnegie Mellon University Qatar
Ossama Obeid, Carnegie Mellon University Qatar
Nizar Habash (advisor), New York University Abu Dhabi

Program Committee Members

Abdelmajid Ben-Hamadou, University of Sfax, Tunisia
Abdelhadi Soudi, Ecole Nationale de l'Industrie Minérale, Morocco
Abdelsalam Nwesri, University of Tripoli, Libya
Achraf Chalabi, Microsoft Research, Egypt
Ahmed Ali, Qatar Computing Research Institute, Qatar
Ahmed Rafea, The American University in Cairo, Egypt
Alexis Nasr, University of Marseille, France
Ali Farghaly, Monterey Peninsula College, USA
Almoataz B. Al-Said, Cairo University, Egypt
Alon Lavie, Carnegie Mellon University, USA

Aly Fahmy, Cairo University, Egypt
Azadeh Shakery, University of Tehran, Iran
Azzeddine Mazroui, University Mohamed I, Morocco
Bassam Haddad, University of Petra, Jordan
Bayan Abu Shawar, Arab Open University, Jordan
Behrang Mohit, Carnegie Mellon University Qatar, Qatar
Eric Atwell, University of Leeds, UK
Farhad Oroumchian, University of Wollongong, Australia
Ghassan Mourad, Université Libanaise, Lebanon
Hassan Sawaf, eBay Inc., USA
Hazem Hajj, American University of Beirut, Lebanon
Hend Alkhalifa, King Saud University, Saudi Arabia
Houda Bouamor, Carnegie Mellon University Qatar, Qatar
Imed Zitouni, Microsoft Research, USA
Joseph Dichy, Université Lyon 2, France
Karim Bouzoubaa , Mohammad V University, Morocco
Karine Megerdooian, The MITRE Corporation, USA
Katrin Kirchhoff, University of Washington, USA
Kemal Oflazer, Carnegie Mellon University Qatar, Qatar
Khaled Shaalan, The British University in Dubai, UAE
Khaled Shaban, Qatar University, Qatar
Khalil Sima'an, Universiteit van Amsterdam, Netherlands
Lamia Hadrach Belguith, University of Sfax, Tunisia
Michael Rosner, University of Malta, Malta
Mohamed Elmahdy, Qatar University, Qatar
Mohsen Rashwan, Cairo University, Egypt
Mona Diab, George Washington University, USA
Mustafa Jarrar, Bir Zeit University, Palestine
Nada Ghneim, Higher Institute for Applied Sciences and Technology, Syria
Nadi Tomeh, University Paris 13, France
Ossama Emam, IBM, USA
Otakar Smr_, D_ám-e D_am Language Institute, Czech Republic
Owen Rambow, Columbia University, USA
Preslav Nakov, Qatar Computing Research Institute, Qatar
Ramzi Abbes, TECHLIMED, France
Salwa Hamada, Cairo University, Egypt
Shahram Khadivi, Tehran Polytechnic, Iran
Sherri Condon , The MITRE Corporation, USA
Taha Zerrouki, University of Bouira, Algeria
Violetta Cavalli-Sforza, Al Akhawayn University, Morocco

Contact

Nizar Habash, New York University Abu Dhabi
habash@ccls.columbia.edu

Paper Submission Instructions

Submission Types

There are two possible submissions types:

- **Full papers** of maximum 8 pages + 2 for references
- **Shared Task System papers** of maximum 4 pages + 2 for references AND authors must have participated in the shared task

Submission Language

Submissions are expected to be in the English language.

Blind Reviewing Policy

The workshop follows a blind reviewing policy. The authors should omit their names and affiliations from the paper and avoid self-references that reveal their identity. **Papers that do not conform to these requirements will be rejected without review.**

Submission Format

All submissions must be electronic in PDF and must be formatted using the EMNLP 2014 [style files](#)

Submission Site

Papers should be submitted via the START Conference Manager site:

<https://www.softconf.com/emnlp2014/arabicnlp14/>

Please do not send papers by email to the organizers. Such papers will not be considered.

Multiple Submission Policy

Papers that have been or will be submitted to other meetings or publications must indicate this at submission time. Authors must inform organizers immediately once a paper is to be withdrawn from the workshop for any reason. Attempting to publish the same paper or with a large overlap (50%) may lead to rejection of the paper even after an acceptance notification have gone out.

Arabic Natural Language Processing Workshop co-located with EMNLP 2014, Doha, Qatar

List of Accepted Papers

We are pleased to announce the accepted papers: 19 workshop papers and 10 shared task papers.

Main Workshop Papers

- Al-Bayan: An Arabic Question Answering System for the Holy Quran
Heba Abdelnasser, Maha Ragab, Reham Mohamed, Alaa Mohamed, Bassant Farouk, Nagwa El-Makky and Marwan Torki
- Annotating corpus data for a constructional analysis of motion verbs in Modern Standard Arabic
Dana Abdulrahim
- Combining strategies for tagging and parsing Arabic
Maytham Alabbas and Allan Ramsay
- The International Corpus of Arabic: Compilation, Analysis and Evaluation
Sameh Alansary
- Uniform Word Segmentation Improves Dialectal Arabic to English Machine Translation
Kamla Al-Mannai, Hassan Sajjad, Alaa Khader, Fahad Al Obaidli, Preslav Nakov and Stephan Vogel
- Automatic Arabic diacritics restoration based on deep nets
Ahmad Al Sallab, Mohsen Rashwan, Hazem M. Raafat and Ahmed Rafea
- A Large Scale Arabic Sentiment Lexicon for Arabic Opinion Mining
Gilbert Badaro, Ramy Baly, Hazem Hajj, Nizar Habash and Wassim El-Hajj
- Transliteration of Arabizi into Arabic Orthography: Developing a Parallel Annotated Arabizi-Arabic Script SMS/Chat Corpus
Ann Bies, Zhiyi Song, Mohamed Maamouri, Stephen Grimes, Haejoong Lee, Jonathan Wright, Stephanie Strassel, Nizar Habash, Ramy Eskander, and Owen Rambow
- Tunisian dialect Wordnet : creation and enrichment using web resources and other Wordnets
Rihab Bouchlaghem, Aymen Elkhilfi and Rim Faiz
- Arabizi Detection and Conversion to Arabic
Kareem Darwish
- A Framework for Creating an MWE Computational Lexical Resource for Dialectal Arabic
Abdelati Hawwari, Mohammed Attia and Mona Diab
- Building a Corpus for Palestinian Arabic: a Preliminary Study
Mustafa Jarrar, Nizar Habash, Diyam Akra and Nasser Zalmout
- Domain and Dialect Adaptation for Machine Translation into Egyptian Arabic
Serena Jeblee, Weston Feely, Houda Bouamor, Alon Lavie, Nizar Habash and Kemal Oflazer
- Semantic Query Expansion for Arabic Information Retrieval
Ashraf Y. Mahgoub, Mohsen A. Rashwan, Hazem Rafea, Mohamed A. Zahran, and Magda B. Fayek
- Arabic Native Language Identification
Shervin Malmasi and Mark Dras

- Using Twitter to Collect a Multi-Dialectal Corpus of Arabic
Hamdy Mubarak and Kareem Darwish
- Evaluating Distant Supervision for Subjectivity and Sentiment Analysis on Arabic Twitter Feeds
Eshrag Refaee and Verena Rieser
- AIDArabic A Named-Entity Disambiguation Framework for Arabic Text
Mohamed Amir Yosef, Marc Spaniol and Gerhard Weikum
- Named Entity Recognition System for Dialectal Arabic
Ayah Zirikly and Mona Diab

Shared Task Papers

- The First shared Task on Automatic Text Correction for Arabic
Behrang Mohit, Alla Rozovskaya, Nizar Habash, Wajdi Zaghouni and Ossama Obeid
- GWU-HASP: Hybrid Arabic Spelling and Punctuation Corrector
Mona Diab, Mohammed Attia and Mohamed Al-Badrashiny
- Arabic Spelling Correction using Supervised Learning
Youssef Hassan, Mohamed Aly and Amir Atiya
- CMUQ@The 2014 Automatic Arabic Error Correction Shared Task
Serena Jeblee, Houda Bouamor, Wajdi Zaghouni and Kemal Oflazer
- TECHLIMED system description for the Shared Task on Automatic Arabic Error Correction
Djamel Mostefa, Omar Asbayou and Ramzi Abbes
- Automatic Correction of Arabic Text: a Cascaded Approach
Hamdy Mubarak and Kareem Darwish
- Fast and Robust Arabic Error Correction System
Michael Nawar and Moheb Ragheb
- The Columbia System in the QALB-2014 Shared Task on Arabic Error Correction
Alla Rozovskaya, Nizar Habash, Ramy Eskander, Noura Farra and Wael Salloum
- A Pipeline Approach to Supervised Automatic Error Correction
Nadi Tomeh, Nizar Habash, Ramy Eskander and Joseph Le Roux
- Auto-correction of arabic common errors for large text corpus
Taha Zerrouki, Khaled Alhowaity and Amar Balla

Session #3 Translation & Transliteration

- Domain and Dialect Adaptation for Machine Translation into Egyptian Arabic
Serena Jeblee, Weston Feely, Houda Bouamor, Alon Lavie, Nizar Habash and Kemal Oflazer
- Uniform Word Segmentation Improves Dialectal Arabic to English Machine Translation
Kamla Al-Mannai, Hassan Sajjad, Alaa Khader, Fahad Al Obaidli, Preslav Nakov and Stephan Vogel
- Arabizi Detection and Conversion to Arabic
Kareem Darwish

Shared Task Main Paper

- The First shared Task on Automatic Text Correction for Arabic
Behrang Mohit, Alla Rozovskaya, Nizar Habash, Wajdi Zaghouni and Ossama Obeid

Poster Session: Main Workshop Papers

- Al-Bayan: An Arabic Question Answering System for the Holy Quran
Heba Abdelnasser, Maha Ragab, Reham Mohamed, Alaa Mohamed, Bassant Farouk, Nagwa El-Makky and Marwan Turki
- Combining strategies for tagging and parsing Arabic
Maytham Alabbas and Allan Ramsay
- Automatic Arabic diacritics restoration based on deep nets
Ahmad Al Sallab, Mohsen Rashwan, Hazem M. Raafat and Ahmed Rafea
- Transliteration of Arabizi into Arabic Orthography: Developing a Parallel Annotated Arabizi-Arabic Script SMS/Chat Corpus
Ann Bies, Zhiyi Song, Mohamed Maamouri, Stephen Grimes, Haejoong Lee, Jonathan Wright, Stephanie Strassel, Nizar Habash, Ramy Eskander, and Owen Rambow
- Tunisian dialect Wordnet : creation and enrichment using web resources and other Wordnets
Rihab Bouchlaghem, Aymen Elkhilfi and Rim Faiz
- A Framework for Creating an MWE Computational Lexical Resource for Dialectal Arabic
Abdelati Hawwari, Mohammed Attia and Mona Diab
- Semantic Query Expansion for Arabic Information Retrieval
Ashraf Y. Mahgoub, Mohsen A. Rashwan, Hazem Rafea, Mohamed A. Zahran, and Magda B. Fayek
- Named Entity Recognition System for Dialectal Arabic
Ayah Zirikly and Mona Diab

Poster Session: Shared Task Papers

- GWU-HASP: Hybrid Arabic Spelling and Punctuation Corrector
Mona Diab, Mohammed Attia and Mohamed Al-Badrashiny
- Arabic Spelling Correction using Supervised Learning
Youssef Hassan, Mohamed Aly and Amir Atiya
- CMUQ@The 2014 Automatic Arabic Error Correction Shared Task
Serena Jeblee, Houda Bouamor, Wajdi Zaghouni and Kemal Oflazer
- TECHLIMED system description for the Shared Task on Automatic Arabic Error Correction
Djamel Mostefa, Omar Asbayou and Ramzi Abbes

- Automatic Correction of Arabic Text: a Cascaded Approach
Hamdy Mubarak and Kareem Darwish
- Fast and Robust Arabic Error Correction System
Michael Nawar and Moheb Ragheb
- The Columbia System in the QALB-2014 Shared Task on Arabic Error Correction
Alla Rozovskaya, Nizar Habash, Ramy Eskander, Noura Farra and Wael Salloum
- A Pipeline Approach to Supervised Automatic Error Correction
Nadi Tomeh, Nizar Habash, Ramy Eskander and Joseph Le Roux
- Auto-correction of Arabic common errors for large text corpus
Taha Zerrouki, Khaled Alhowaity and Amar Balla