
THE CONFERENCE PROGRAMME AT A GLANCE

APRIL 22, 2009

ROOM: FARHATY 1 & 2

9:00-9:30	Opening Ceremony
9:30-10:15	Keynote Speech: Why is SMT so hard to understand? By Prof. Andy Way
10:15-11:15	O1 - Text Categorisation, Summarisation, Applications
10:15-10:35	Arabic Keyphrase Extraction using Linguistic Knowledge and Machine Learning Techniques, Tarek El-Shishtawy & Abdulwahab Al-Sammak
10:35-10:55	Using EM for Text Classification on Arabic Documents, Ghassan Kanaan, Mustafa Yaseen, Riyad Al-Shalabi, Bashar Al-Sarayreh & Ashraf Bany Mustafa
10:55-11:15	Assessing Word-form Based Search for Information in Arabic: Towards a new Type of Lexical Resource, Mouna Anizi & Joseph Dichy
11:15-11:45	Break
11:45-12:45	P1 - Morphology, Tagging, Diacritization
	Word Stress and Vowel Neutralization in Modern Standard Arabic, Jack Halpern
	Bootstrapping Tagged Islamic Corpora, Mahmoud Shokrollahi-Far, Behrooz Minaei, Issa Barzegar, Hadi Hossein-Zade, Mojde Ghasdi & Seyed-salman Hoseini
	A Hybrid System for Automatic Arabic Diacritization, Mohsen Rashwan, Mohammad Al-Badrashiny, Mohammed Attia & Sherif Abdou
	Implementation of Infixes and Circumfixes in the Spellcheckers, Taha Zerrouki & Amat Balla
	Arabic Part-Of-Speech Tagging using Transformation-Based Learning, Shabib AlGahtani, William Black & John McNaught
	Classification of Arabic Information Extraction Methods, Abd El Salam Alhajjar, Mohammad Hajjar & Khaldoun Zreik
	Processing Large Arabic Text Corpora: Preliminary Analysis and Results, Fahad Alotaiby, Ibrahim Alkharashi & Salah Foda
	Enhancing the ElixirFM Lexicon with Verbal Valency Frames, Viktor Bielicky & Otakar Smrz
12:45-14:00	Lunch
14:00-15:40	O2 - Handwritten characters
14:00-14:20	A new Approach for Online Arabic Handwriting Recognition, Monji Kherallah, Lobna Haddad & Adel M. Alimi
14:20-14:40	Artificial Tutor For Arabic Handwriting Training, Sherif Abdou & Aly Fahmy
14:40-15:00	Arabic Word and Text Recognition - Current Developments, Volker Märgner & Haikal El Abed
15:00-15:20	Linguistic Resources for Arabic Handwriting Recognition, Stephanie Strassel
15:20-15:40	A Multi Agent System for Recognizing Printed Arabic Words, Mohamed Ben Halima & Adel M. Alimi
15:40-16:10	Break
16:10-17:10	P2 - Arabic Toolkits, Infrastructures
	Linguistic Constraints as a Sub-component of an Arabic Linguistic Infrastructure, Salwa Hamada & Adil Al Kufaishi
	Collaborative Arabic Morphological Service, Natheer Khasawneh, Eyad Taqieddin & Ahmad Al-Hammouri
	MADA+TOKAN: A Toolkit for Arabic Tokenization, Diacritization, Morphological Disambiguation, POS Tagging, Stemming and Lemmatization, Nizar Habash, Owen Rambow & Ryan Roth
	Minimal Resources for Arabic Parsing: an Interactive Method for the Construction of Evolutive Automata, Claude Audebert, Christian Gaubert & André Jaccarini
	High Performance Multilingual Natural Language Processing Framework for Industrial Applications, Amine Akrouf, D.J McCloskey, Badr AlDaihani & Shabib AlGahtani
	An RSS Feed Analysis Application and Corpus Builder, Shereen Khoja
	Collaborative Construction of Arabic Lexical Resources, Mohammad Daoud, Daoud Daoud & Christian Boitet
16:10-17:10	P3 - Applications and Evaluation of Applications
	In the Heart of Semantic Technology, Paolo Lombardi
	Evaluation Approaches for an Arabic Extractive Generic Text Summarization System, Ibrahim Sobh, Nevin Darwish & Magda Fayek
	Arabic Query Expansion Using Interactive Word Sense Disambiguation, Riyad Al-Shalabi, Ghassan Kanaan, Mustafa Yaseen, Bashar Al-Sarayreh & Nada A. Al-Naji
	Addaall Arabic Search Engine: Improving Search based on Combination of Morphological Analysis and Generation Considering Semantic Patterns, Mamoun Hattab, Bassam Haddad, Mustafa Yaseen, Asem Duraidi & Abdullah Abu Shmais
	Arabic, English and French: Three Languages in a Filtering Systems Evaluation Project, Romaric Besançon, Djamel Mostefa, Ismail Timimi, Stéphane Chaudiron, Mariama Laïb & Khalid Choukri
17:10-18:30	O3 - Treebank
17:10-17:30	Syntactic Annotation in the Columbia Arabic Treebank, Nizar Habash, Reem Faraj & Ryan Roth
17:30-17:50	A Statistical Method for Detecting the Arabic Empty Category, Hitham M. Abo Bakr, Khaled Shaalan & Ibrahim Ziedan
17:50-18:10	A Tool for Syntactic Tagging an Arabic Treebank, Fériel Ben Fraj, Chiraz Ben Othmane Zribi & Mohamed Ben Ahmed
18:10-18:30	Creating a Methodology for Large-Scale Correction of Treebank Annotation: the Case of the Arabic Treebank, Mohamed Maamouri, Ann Bies & Seth Kulick

THE CONFERENCE PROGRAMME AT A GLANCE

APRIL 23, 2009

ROOM: FARHATY 1 & 2

9:00-10:20

O4 - Large Projects and Initiatives

- 9:00-9:20 **MEDAR: Arabic Language Technology, State-of-the-art and a Cooperation Roadmap**, *Bente Maegaard, Mohammed Attia, Khalid Choukri, Steven Krauwer, Chafic Mokbel & Mustafa Yaseen*
- 9:20-9:40 **Arabic Language Resources and Tools for Speech and Natural Language: KACST and Balamand**, *Mansour Alghamdi, Chafic Mokbel & Mohamed Mrayati*
- 9:40-10:00 **Arabic Language Resources in HIAST**, *Oumayma Al Dakkak, Nada Ghneim, Afaf Alshalaby, Riad Sonbol & Mhd. Saïd Desouki*
- 10:00-10:20 **MEDAR: Mediterranean Arabic Language and Speech Technology Inventory of the HLT Products, Players, Projects, and Language Resources**, *Khalid Choukri*

10:20-11:20

P4 - Named Entities and Related Issues

- Lexicon-Driven Approach to the Recognition of Arabic Named Entities**, *Jack Halpern*
- Transliteration using Phrase-based SMT Approach on Substrings**, *Sara Noeman*
- Person Named Entity Generation and Recognition for Arabic Language**, *Ibrahim Alkharashi*
- An Optimized Method For Arabic Cross Document Named Entity Normalization**, *Khaled S. Refaat & Amgad Madkour*
- A Multilingual Named Entities Corpus for Arabic, English and French**, *Djamel Mostefa, Mariama Laïb, Stéphane Chaudiron, Khalid Choukri & Gaël de Chalendar*

10:20-11:20

P5 – Machine Translation and Related

- REMOOV: A Tool for Online Handling of Out-of-Vocabulary Words in Machine Translation**, *Nizar Habash*
- Designing an XML Lexicon Architecture for Arabic Machine Translation Based on Role and Reference Grammar**, *Yasser Salem & Brian Nolan*

11:20-11:50

Break

11:50-12:50

P6 - Text Categorisation, Classifications

- A Study of Arabic Text Preprocessing Methods for Text Categorization**, *Dina Said, Nayer Wanas, Nevin Darwish & Nadia Hegazy*
- Towards a Classification System for Arabic Texts**, *Rami Ayadi, Mohsen Maraoui & Mounir Zrigui*
- A Trigger-based Classifier**, *Mourad Abbas, Kamel Smaili & Daoud Berkani*

11:50-12:50

P7 - Speech Resources & Technologies

- An HMM System for Recognizing Articulation Features for Arabic Phones**, *Hosam Hammady, Sherif Abdou, Mostafa Shahin, Mohsen Rashwan & Ossama Badawy*
- Stem-based Arabic Language Models Experiments**, *Mohsen Moftah, Waleed Fakhr, Sherif Abdou & Mohsen Rashwan*
- A Large Arabic Broadcast News Speech Data Collection**, *Niklas Paulsson, Khalid Choukri, Djamel Mostefa, Denise DiPersio, Meghan Glenn & Stephanie Strassel*
- Second Generation Tools (AMIRA 2.0): Fast and Robust Tokenization, POS Tagging, and Base Phrase Chunking**, *Mona Diab*

12:50-14:00

Lunch

14:00-15:40

Roadmap: The MEDAR cooperation roadmap for Arabic language technology.

15:40-16:10

Break

16:10-17:50

O5 - Morphology, Tagging, Diacritization

- 16:10-16:30 **Arabic Part-Of-Speech Tagging using the Sentence Structure**, *Yahya Ould Mohamed El Hadj, Imad Abdulrahman Al-Sughayeir & Abdullah Mahdi Al-Ansari*
- 16:30-16:50 **A Framework for the Rapid Development of List Based Domain Specific Arabic Stemmers**, *Samhaa R. El-Beltagy & Ahmed Rafea*
- 16:50-17:10 **Sarrif! - The Elegant Arabic Morphology Parser**, *Suhel Jaber & Rodolfo Delmonte*
- 17:10-17:30 **Diacritization and Transliteration of Proper Nouns from Arabic to English**, *Hamdy S. Mubarak, Mohamed Al Sharqawy & Esraa Al Masry*
- 17:30-17:50 **Towards an Automatic Conversion Approach of Editorial Arabic Dictionaries into LMF-ISO 24613 Standardized Model**, *Aida Khemakhem, Imen Elleuch, Bilel Gargouri & Abdelmajid Ben Hamadou*

17:50

Closing Ceremony