NEMLAR Newsletter

News on Arabic language resources and Arabic language technologies

Issue 15

December 2008

Newsletter content:

1. 2nd International Conference on Arabic Resources and Tools
2. Other upcoming events
3. Join the Network
4. How to contribute to the Network

1. 2nd International Conference on Arabic Resources and Tools

The conference is organized by the MEDAR consortium.
It will take place the **22-23 April 2009, in Cairo, Egypt.**

See the conference web site: http://www.medar.info/conference/index.php

Conference aims

Language Resources (LRs) are a central component of the linguistic infrastructure, necessary for the development of HLT applications and products, and therefore for industrial development. In this conference we will focus on Arabic language technology and on the necessary language resources and tools for both research and commercial development of language technology for Arabic.
Multilingual language technology is a particular focus, as well as general methodologies. The other important aspect for the promotion of Arabic language technology is cooperation: Cooperation is extremely important for the advancement of the field, be it cooperation between European, Arabic and American partners, cooperation between Arabic partners, cooperation between research and industry etc. MEDAR will present its first proposal for a **Cooperation Roadmap**, and will seek discussion and collaboration for its final version.

The aim of this conference is to provide a forum for the discussion of the state-of-the-art for Arabic resources and tools, in particular for machine translation and multilingual information retrieval, discuss problems and opportunities, exchange information regarding LRs and tools, their applications, ongoing and planned activities, industrial uses and needs, requirements coming from the new e-society, both with respect to policy issues and to technological and organisational ones. The roadmap discussions will in particular bring in policy issues and collaboration and organisation issues.

Conference topics

Examples of the topics which may be addressed by papers submitted to the conference are given below. The conference programme will have a particular focus on the multilinguality issues mentioned above, but will not be restricted to these.

Issues in the design, construction and use of Arabic Language Resources (LRs) and tools

- Guidelines, standards, specifications, models and best practices for Arabic LRs;
- Methods, tools and procedures for the acquisition, creation, management, access, distribution and use of Arabic LRs;
- Methods for the extraction and acquisition of knowledge (e.g. terms, lexical information, language modelling) from LRs;
- Organisational and legal issues in the construction, distribution, access and use of LRs;
- Availability and use of generic vs. task/domain specific LRs;
- Definition and requirements for a Basic LAnguage Resource Kit (BLARK) for Arabic;
- Monolingual and multilingual LRs;
- Multimedia and multimodal LRs. - Integration of various media and modalities in LRs (speech, vision, language);
- Documentation and archiving, including minority and endangered languages;
- Ontologies and knowledge representation;
- Terminology, term extraction, domain-specific dictionaries;
- LRs for linguistic research in human-machine communication;
- Exploitation of LRs in different types of applications (information extraction, information retrieval, speech dictation, translation, summarisation, web services, semantic web, etc.);
- Exploitation of LRs in different types of interfaces (dialog systems, natural language and multimodal/multisensorial interactions, etc.);
- Industrial LRs requirements, user needs and community's response;
- Industrial use of LRs;
- Metadata descriptions of LRs.
- Taggers
- Parsers
- Alignment tools
  Etc.

Issues in multilingual Arabic language technology

- Machine translation to or from Arabic
- Multilingual information retrieval
- Multilingual document retrieval
- Spoken translation
- Multilingual spoken document retrieval
  Etc.

Issues in Human Language Technologies (HLT) evaluation
- Evaluation, validation, quality assurance of Arabic LRs,
- Evaluation methodologies, protocols and measures,
- Usability and user experience evaluation, qualitative and perceptive evaluation,
- Evaluation in multilingual Arabic language processing (machine translation, multilingual information retrieval, cross-lingual information retrieval etc.)
- From evaluation to standardisation.

General issues
- National and international activities and projects on Arabic
- Roadmapping for Arabic language technology
- Needs, possibilities, forms, initiatives of/for international cooperation, and their organisational and technological implications
- Open architectures for LRs and tools
- Arabic LRs and the needs/opportunities of the emerging industries
- Arabic LRs and contributions to societal needs (e.g. e-society)

Programme
The Scientific Programme will include invited talks, oral and poster presentations, referenced demonstrations and panels.

Abstract submission
The conference will be using an electronic submission tool. It will soon be available at the conference website, www.medar.info/conference.

Important dates
- **30 January 2009:** Submission of proposals for papers, posters, referenced demos
- **9 March 2009:** Notification of acceptance
- **6 April 2009:** Final versions for the proceedings

   NB. There will be no extension of deadlines!

The proceedings of the conference will include both normal papers, poster papers and papers that go with a demonstration. Proceedings will be electronic, but the abstracts will be printed.

Internet connections and various computer platforms and facilities will be available at the conference site. In addition to referenced demos concerning LRs and related tools, it will be possible to run unreferenced demos of language processing products, systems and tools. You should contact us (details posted on www.medar.info) if this opportunity interests you.

Tutorials

Tutorials will be organised before the conference. Please check the conference website.

Conference programme committee
- Bente Maegaard, CST, University of Copenhagen, Denmark (co-chair)
- Khalid Choukri, ELDA, Paris, France (co-chair)
- Mohsen Rashwan, Cairo University, Egypt
- Ossama Emam, IBM Egypt
- Mustafa Yaseen, Amman University, Jordan
- Chafik Mokbel, University of Balamand, Tripoli, Lebanon
- Abdelhak Mouradi, ENSIAS, University of Mohammed V Soussi, Rabat, Morocco
- Mansour Al-Ghamdi, King Abdulaziz City for Science and Technology, Saudi Arabia (tbc)
- Ibrahim Al-Kharashi, King Abdulaziz City for Science and Technology, Saudi Arabia
- Steven Krauwer, Utrecht University, The Netherlands
- Mona Diab, Columbia University, USA
- Nizar Habash, Columbia University, USA
- Chris Cieri, LDC, USA

Organising Committee
- Mohsen Rashwan, Cairo University, Egypt
- Mohamed Attia, RDI, Cairo, Egypt
- Hélène Mazo, ELDA, France
- Dorte Haltrup Hansen, CST, University of Copenhagen

2. Upcoming Events
- ➢ 3rd International Conference on Arabic Language Processing (CITALA'09)
  May 4th - 5th 2009,
  Rabat, Morocco.
  See more about the event

- ➢ (**EAMT 2009**)
  13th Annual Conference of the European Association for Machine Translation
  May 14-15, 2009,
  Universitat Politècnica de Catalunya, Barcelona, SPAIN

3. Join the Network

NEMLAR started a network for players who want to collaborate for the advancement of Arabic language technology. This network is being continued by MEDAR.

All institutions and enterprises are invited to join the network, if you share our mission. By joining you will get access to the newest information about language technology (actors, products, research) for Arabic in the region, you will be able to participate in the surveys and in the creation of the roadmap for cooperation. You may also find new partners and collaborators for your future activities.

Join the network through the website www.medar.info, and receive the newsletter and other information.

4. How to contribute to the Network

Please help by sending us information on coming events, new software, resources, books, papers and journals. We will distribute it via this newsletter and put it on the MEDAR homepage www.medar.info .

If you wish to contribute to the work of the MEDAR project you please fill in the survey questionnaire about your language resources and/or the industry needs for language resources. You can find a link to the survey questionnaire at: www.medar.info

To subscribe or unsubscribe to the newsletter, please send an email to: nemlar@hum.ku.dk

The newsletter is published by the MEDAR project www.medar.info and produced by Centre for Language Technology, University of Copenhagen, Denmark.

Contact: Bente Maegaard (Coordinator of MEDAR) or
             Dorte Haltrup Hansen,
             Centre for Language Technology,
             University of Copenhagen,
             Njalsgade 140 – 142,
             Copenhagen S,
             Denmark
             Tel: +45 35 32 90 74, Fax: +45 35 32 90 89
             nemlar@hum.ku.dk