Special Session on Arabic Language Processing

The Arabic language is considered to be one of the most difficult languages for written and spoken language processing due to its morphological, syntactic, phonetic and phonologic properties. It is also one of the six official languages of the United Nations. The rapid growth of the Internet has significantly facilitated the dissemination of electronic Arabic documents in recent years. As a result, Arabic Natural Language Processing (ANLP) has moved from the small academic research labs to the main stream of the software industry research and development labs.

Research on written Arabic language processing started in the 1970s, even before the problems of Arabic text editing were completely solved. The first studies focused primarily on lexicons and morphology.

Current research and development in Arabic Language Processing spans a wide range of application areas. This special session is intended to present the state of the art and practice on Arabic Natural Language Processing. The aim of this session is to gather and reinforce collaboration between researchers from both the written and spoken Arabic language processing communities. It will also offer the opportunity to discuss recent advances on both the scientific and application sides of the problem, in monolingual and multilingual contexts.

This special session on written and spoken Arabic processing includes (but is not limited to) the following topics:

- Cross Language Information Retrieval
- Arabic search engines and page ranking
- Machine readable dictionary and corpora
- Machine translation
- Web crawling and indexes for Arabic pages
- Disambiguation and Part of speech tagging
- Text classification
- Semantic analysis
- Morphological analysis
- Text parsing and generation,
- Discourse analysis
- Text summarization
- Flat and hierarchical clustering
- Tolerant retrieval and phonetic correction
- Arabic characters recognition
- Speech recognition
- Text to speech synthesis,
- Automatic speaker and language identification

Session Chair:
Dr. Ahmad T. Al-Taani, Yarmouk University, Jordan.

For submission:
and select: Natural Language Processing (Arabic) topic