Roadmap for Language Resources and Evaluation In a Multilingual Environment:

Organised by COCOSDA and WRITE

Genoa, 28 May 2006
In Conjunction with LREC 2006

Following the successful Joint COCOSDA and WRITE Meeting, held at LREC 2004 in, Lisbon on **Building the Language Resources and Evaluation Roadmap,** our focus of the 2006 workshop will be on the challenges for language resources in a multilingual environment. This will include, as a consequence, also language, speech and multimodal technologies and applications (based on LRs and evaluation), and infrastructural issues (such as standards, distributed resources and web services, international cooperation and funding policies).

Invited presenters include COCOSDA and WRITE members, organizers of relevant LREC panels and workshops, as well as representatives of funding agencies.

Presentations will be focused on emerging common trends and priorities for next years, also exploiting results of other LREC events, to try to form a synthesis and use it for a vision towards the future. Discussion will be focused on where the whole field is moving and which initiatives are needed.

Important topics for discussion include:

1) How to re-direct the field of language resources at an important juncture of the field. The "classic" language resources work involving corpora, lexica, and wordnets etc. has matured to the extent that international standards are being stipulated. However, we face new challenges such as mutlilingualism, mutli-modality, as well as knoweldge-based processing.

2) The emphasis on mutlingualism and mutli-modality underlines the fact that the field of language resources cannot grow effectively without international and cross-disciplinary collaboration.

A purpose of this workshop is to gather leading scholars in this field to generate synergy among their visions, and to formulate recommendations to funding agencies based on a synergized vision.

By joining the efforts of COCOSDA and WRITE, as well as of other LREC workshops, another important goal is to facilitate an infrastructure for multi-national and multi-disciplinary collaboration following the above-mentioned recommendations.

Possible other topics with a potential impact:
- How to turn LRs into valuable knowledge content and/or societal recognition
- How to navigate geo-political and infrastructural barriers to access distributed LRs
- How to fully explore the commercial and technical advantages of highly computerised languages and address the needs of less computerised languages at the same time
- How to develop uniform and implementable LR standards for speech, text and multimodal applications without sacrificing language diversity and language identity
- How to incorporate language and cultural specific literary (written and spoken) traditions into the electronic language resources and create synergy

Confirmed presenters include: Nick Campbell, Dafydd Gibbon, Nancy Ide, Noriko Kando, Steven Krauwer, Joseph Mariani, Jean-Claude Martin, Stelios Piperidis, Justus Roux, Peter Wittenburg, Pierre Zweigenbaum.

Time schedule

- 9.30-11.00: session
-11.00-11.30: coffee
-11.30-13.30: session
-13.30-14.30: lunch
-14.30-16.30: session
-16.30-17.00: coffee & end

Organizing Committee

Nicoletta Calzolari (ILC-CNR, Pisa – Italy)

Khalid Choukri (ELDA, Paris – France)

Christopher Cieri (University of Pennsylvania – Linguistic Data Consortium, Philadelphia – USA)

Dafydd Gibbon (University of Bielefeld, Bielefeld – Germany)

Chu-Ren Huang (Academia Sinica, Taipei – Taiwan)

For:

COCOSDA: International *Co*mmittee for *Co*-ordination and *S*tandardisation of Speech *D*atabases and *A*ssessment Techniques

WRITE: International Committee for Written Resources Infrastructure, Technology, and Evaluation