LREC 2004
24-30 May 2004, Lisbon, Portugal

DEADLINE EXTENSION FOR WORKSHOP AND PANEL PROPOSALS
Deadline extended to 27th October.

ABSTRACTS SUBMISSION FORMS NOW AVAILABLE ON-LINE:

http://www.lrec-conf.org

The fourth international conference on Language Resources and Evaluation, LREC 2004, is organised by ELRA in cooperation with other Associations and Consortia, including ACL, AFNLP, ALLC, ALTA, COCOSDA and Oriental COCOSDA, EAFT, EAMT, ELSNET, ENABLER, EURALEX, GKS, GWA, IAMT, ICWLR, ISCA, LDC, ONTOWEB, TEI, and with major national and international organisations, including the Commission of the EU - Information Society DG, Unit E1 "Interfaces and Cognition". Co-operation with other organisations is currently being sought.

Location

Centro Cultural de Belem, Lisbon, Portugal

Dates

- Pre-conference workshops: 24-25 May 2004
- Main conference: 26-27-28 May 2004
- Post-conference workshops: 29-30 May 2004

Conference web site: http://www.lrec-conf.org

Conference aims

In the Information Society, the pervasive character of Human Language Technologies (HLT) and their relevance to practically all fields of Information Society Technologies (IST) has been widely recognised.

Two issues are particularly relevant: the availability of Language Resources (LRs) and the methods for the evaluation of resources, technologies, products and applications. Substantial mutual benefits are achieved by addressing these issues through international collaboration.

The term "language resources" (LRs) refers to sets of language data and descriptions in machine readable form, used in many types of areas/components/systems/applications:
- Creation and evaluation of natural language, speech and multimodal algorithms and systems;
- Software localisation and language services;
- Language enabled information and communication services;
- Knowledge management;
- E-commerce, e-publishing, e-learning, e-government;
- Cultural heritage;
- Linguistic studies;
- Etc.

This large range of uses makes the LRs infrastructure a strategic part of the e-society, where the creation of a basic set of LRs for all languages must be ensured in order to bring all languages to the same level of usability and availability.

Examples of LRs are written or spoken corpora and lexica, which may be annotated or not, multimodal resources, grammars, terminology or domain specific databases and dictionaries, ontologies, multimedia databases, etc. LRs also cover basic software tools for the acquisition, preparation, collection, management, customisation and use of the above mentioned examples.

The relevance of evaluation for language technologies development is increasingly recognised. This involves assessing the state-of-the-art for a given technology, measuring the progress achieved within a programme, comparing different approaches to a given problem, assessing the availability of technologies for a given application, benchmarking, and assessing system usability and user satisfaction.

The aim of this conference is to provide an overview of the state-of-the-art, discuss problems and opportunities, exchange information regarding LRs, 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. LREC will also elaborate on evaluation methodologies and tools, explore the different trends and promote initiatives for international collaboration in the areas mentioned above.

Conference topics

Examples of the topics which may be addressed by papers submitted to the conference are given below.

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

- Guidelines, standards, specifications, models and best practices for LRs;
- Methods, tools and procedures for the acquisition, creation, management, access, distribution and use of 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 and Extended LAnguage Resource Kit (BLARK, ELARK) for all languages;
- Monolingual and multilingual LRs;
- Multimedia and multimodal LRs. - Integration of various media and modalities in LRs (speech, vision, language);
- Documentation and archiving of languages, including minority and endangered languages;
- Ontologies and knowledge representation ;
- Terminology and NLP, tools and methodologies for terminology and ontology building, term extraction, specialised 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 production of LRs;
- Industrial use of LRs;
- Metadata descriptions of LRs.

Issues in Human Language Technologies (HLT) evaluation

- Evaluation, validation, quality assurance of LRs,
- Evaluation methodologies, protocols and measures,
- Benchmarking of systems and products, resources for benchmarking and evaluation, blackbox, glassbox and diagnostic evaluation of systems,
- Usability and user experience evaluation, qualitative and perceptive evaluation,
- Evaluation in written language processing (document production and management, text retrieval, terminology extraction, message understanding, text alignment, machine translation, morphosyntactic tagging, parsing, semantic tagging, word sense disambiguation, text understanding, summarisation, question answering, localisation, etc.);
- Evaluation in spoken language processing (speech recognition and understanding, voice dictation, oral dialog, speech synthesis, speech coding, speaker and language recognition, spoken translation, etc.);
- Evaluation of multimedia document retrieval and search systems (including detection, indexing, filtering, alert, question answering, etc);
- Evaluation of multimodal systems;
- From evaluation to standardisation.

General issues

- National and international activities and projects;
- LRs and the needs/opportunities of the emerging industries;
- LRs and contributions to societal needs (e.g. e-society);
- Priorities, perspectives, strategies in national and international policies for LRs;
- Needs, possibilities, forms, initiatives of/for international cooperation, and their organisational and technological implications;
- Open architectures for LRs.

The Conference targets the integration of different types of LRs (spoken, written and other modalities) and of the respective communities. To this end, LREC encourages submissions covering issues which are common to different types of Language Technologies, such as dialog strategy, written and spoken translation, domain-specific data, multimodal communication or multimedia document processing, and will organise, in addition to the usual tracks, common sessions encompassing the different areas of LRs.

Programme

The Scientific Programme will include invited talks, oral presentations, poster presentations, referenced demonstrations and panels. There is no difference in quality between oral presentations and poster presentations. Only the appropriateness of the type of communication to the content of the paper will be considered.

Abstract submission

On-line submission forms are now available.

On the LREC 2004 web pages, http://www.lrec-conf.org/lrec2004/index.php, from the "Abstract submission" section, you can choose the appropriate submission form and submit paper, demonstration or poster abstracts.
A submission form is also available on the LREC 2004 web site if you would like to propose a panel.

Workshops

Submission forms for workshops' proposals are also available on-line.

Pre-conference workshops will be organised on 24th and 25th May 2004, and post-conference workshops on 29th and 30th May 2004. A workshop is normally either half day or full day.

The workshop proposers will be responsible for the organisational aspects (e.g. workshop call preparation and distribution, review of papers, notification of acceptance, assembling of the workshop proceedings, etc.).

Proceedings will be printed for each workshop.

Important dates

- Submission of proposals for panels and workshops: 27th October 2003
- Submission of proposals for oral and poster papers, referenced demos: 31st October 2003
- Notification of acceptance of workshop and panel proposals: 14th November 2003
- Notification of acceptance of oral papers, posters, referenced demos: 23rd January 2004
- Final versions for the proceedings: 1st March 2004

- Conference: 26th, 27th and 28th May 2004
- Pre-conference workshops: 24th and 25th May 2004
- Post-conference workshops: 29th and 30th May 2004

The proceedings of the conference will include both oral and poster papers.

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.lrec-conf.org) if this opportunity interests you.

Consortia and project meetings

Consortia or projects wishing to take this opportunity for organising meetings should contact the ELDA office, lrec@elda.fr (further details are given at the end of the document).

Conference registration fees

	Before 20th February 2004	After 20th February 2004	On-site
Standard participant	260 Euro	320 Euro	380 Euro
Participant from ELRA member organisation	210 Euro	250 Euro	290 Euro
Student	110 Euro	130 Euro	150 Euro

The fees cover the following services: a copy of the proceedings, welcome reception, conference dinner, coffee-breaks and refreshments.

Workshop registration fees

	½ day	Full day
Workshop and conference participant	50 Euro	100 Euro
Workshop only participant	85 Euro	170 Euro

The fees cover the following services: a copy of the proceedings of the attended workshop, coffee-breaks and refreshments.

Conference programme committee

- Nicoletta Calzolari, Istituto di Linguistica Computazionale del CNR, Pisa, Italy (Conference Chair)
- Khalid Choukri, ELRA, Paris, France
- Teresa Lino, Universidade Nova de Lisboa, Lisbon, Portugal
- Bente Maegaard, CST, Copenhagen, Denmark
- Joseph Mariani, LIMSI-CNRS, Orsay, France
- Jan Odijk, UIL-OTS, Utrecht, the Netherlands, and ScanSoft, Merelbeke, Belgium
- Daniel Tapias, Telefonica Moviles, Madrid, Spain

The composition of the committees as well as instructions and addresses for registration and accommodation will be detailed on the LREC web site at www.lrec-conf.org.

ELRA

For more information about ELRA (European Language Resources Association), please contact:

Khalid Choukri, ELRA CEO
55-57 Rue Brillat-Savarin,
75013 Paris - France
Tel: + 33 1 43 13 33 33
Fax: + 33 1 43 13 33 30
Email: choukri@elda.fr
Web: www.elra.info or www.elda.fr

****
The first LREC was organised in Granada (Spain) in 1998: 197 papers and posters were presented, with about 510 registered participants from 38 different countries from all continents. Among these, the largest group came from Spain (81 participants), followed by France (75), USA (73), Germany (47), UK (43) and Italy (41). Registered participants belonged to over 325 different organisations.

LREC 2000, in Athens, had 129 oral papers and 152 posters presented, with around 600 participants from 51 different countries from all continents. Among these, the largest group came from Greece (117), followed by USA (70), France (59), Germany (45), UK (43), Japan (35) and Italy (29). Registered participants belonged to 319 different organisations.

LREC 2002, which took place in Las Palmas de Gran Canaria (Spain), attracted over 700 representatives, coming from 38 countries around the world, who could take advantage of the numerous oral and poster presentations (around 365, covering every area in HLT).

A similar number of participants is expected at LREC 2004.
****