****

CALL FOR PAPERS

LREC 2008 Workshop on

Language Technology for Cultural Heritage Data
(LaTeCH 2008)

Special Theme:
"Resources and Tools for Studying Language
Variety and Change"

1 June, 2008, Marrakech, Morocco

http://ilk.uvt.nl/latech08/

Submission deadline: 20 February 2008

****

The Second Workshop on Language Technology for Cultural Heritage Data
(LaTeCH 2008) will be held in conjunction with LREC 2008, and will
take place on June 1 in Marrakech, Morocco.

* SCOPE AND TOPICS

Museums, archives, and libraries around the world maintain large
collections of cultural and scientific heritage objects, such as
archaeological artefacts, audio and video recordings, or manuscripts,
archival documents and other written sources. Such collections are a
potentially very valuable resource for specialists and laypersons
alike, provided they can be easily accessed and automatically
processed. Furthermore, textual cultural heritage resources, such as
old manuscripts and early printed books, are not only interesting for
their information content, but are also an invaluable source for
linguistic research on diachronic and synchronic language variety and
change. While several large scale digitisation projects are currently
underway to make cultural heritage resources more accessible, it is
equally important to develop powerful tools to search, link, enrich,
and mine the digitised data. Language technology has a crucial role to
play in this, even for collections which are primarily non-textual,
since text is the pervasive medium used for meta-data. At the same
time, the cultural heritage domain poses special challenges for the
NLP community, including the use of historic or non-standard language,
the presence of OCR or transcription errors in the data, and the
necessity to deal with data from various media.

For LaTeCH 2008, we invite papers on language technology for cultural heritage data in general and on the special theme of "Studying Language Variety and Change". Topics of interest include, but are not limited to, the following:

- enriching cultural heritage data by inducing meta-data
- dealing with linguistic variation and non-standard or non-contemporary use of language
- automatic error detection and cleaning
- adapting existing NLP tools for the cultural heritage domain
- linking and retrieving information from different sources, media, and languages
- representing cultural heritage data to different audiences (personalisation, text simplification, text summarisation, text generation from databases, hypertext generation)
- knowledge discovery in cultural heritage data
- complex annotation tools
- determination of word and sentence boundaries within manuscripts
- resources for and treatment of dialects (general solutions)
- annotations of language variety on the orthographic, morphological, and syntactic level
- global language resource management systems
- repositories of cultural and scientific heritage data

* SUBMISSIONS

Authors are invited to submit full papers on original, unpublished work in the topic area of this workshop. Submissions should should not exceed 8 pages and should be typeset using a font size of 11 points. (Style files will be made available by LREC for the camera-ready versions of accepted papers.)

The reviewing of the papers will be blind and the papers should not include the authors' names and affiliations. Each submission will be reviewed by at least two members of the program committee. Accepted papers will be published in the workshop proceedings.

Papers should be submitted electronically, no later than February 20, 2008. The only accepted format for submitted papers is Adobe PDF. Details about the submission procedure will be published on the workshop webpage (http://ilk.uvt.nl/latech08/) closer to the time.

* IMPORTANT DATES

20 February - Deadline for workshop papers
21 March   - Notification of acceptance
 4 April   - Camera-ready papers due
 1 June   - Workshop held at LREC 2008

* PROGRAM COMMITTEE (confirmed)

Ion Androutsopoulos, Athens University of Economics and Business, Greece
Timothy Baldwin, University of Melbourne, Australia
David Bamman, Perseus, USA
David Birnbaum, University of Pittsburgh, USA
Antal van den Bosch, Tilburg University, The Netherlands
Andrea Bozzi, ILC-CNR, Pisa, Italy
Kate Byrne, University of Edinburgh, Scotland
Paul Clough, Sheffield University, UK
Greg Crane, Perseus, USA
Milena P. Dobreva, HATII, University of Glasgow, Scotland
Mick O'Donnell, Universidad Autonoma de Madrid, Spain
Matthew James Driscoll, Kobenhavns Universitet, Denmark
Franciska de Jong, University of Twente, The Netherlands
Claire Grover, University of Edinburgh, Scotland
Ben Hachey, University of Edinburgh, Scotland
Djoerd Hiemstra, University of Twente, The Netherlands
Dolores Iorizzo, Imperial College London, UK
Christer Johansson, University of Bergen, Norway
Piroska Lendvai, Tilburg University, The Netherlands
Anke Luedeling, Humboldt-Universitat, Germany
Roland Meyer, University of Regensburg, Germany
Maria Milosavljevic, University of Edinburgh, Scotland
Marie-Francine Moens, Katholieke Universiteit Leuven, Belgium
Marco Passarotti, Universita Cattolica del Sacro Cuore, Italy
Martin Reynaert, Tilburg University, The Netherlands
Kiril Ribarov, Charles University, Czech Republic
Maarten de Rijke, University of Amsterdam, The Netherlands
Peter Robinson, ITSEE, UK
Maria Simi, University of Pisa, Italy
Caroline Sporleder, Saarland University, Germany

* ORGANISING COMMITTEE

Caroline Sporleder (Co-Chair), Saarland University, Germany
Kiril Ribarov (Co-Chair), Charles University, Czech Republic
Antal van den Bosch, Tilburg University, The Netherlands
Milena P. Dobreva, HATII, University of Glasgow, Scotland
Matthew James Driscoll, Kobenhavns Universitet, Denmark
Claire Grover, University of Edinburgh, Scotland
Piroska Lendvai, Tilburg University, The Netherlands
Anke Luedeling, Humboldt-Universitat, Germany
Marco Passarotti, Universita Cattolica del Sacro Cuore, Italy

* FURTHER INFORMATION

Workshop web page: http://ilk.uvt.nl/latech08/

LREC 2008 web page: http://www.lrec-conf.org/lrec2008/

* CONTACT INFORMATION

Caroline Sporleder, csporled (at) coli.uni-sb.de
Kiril Ribarov, ribarov (at) ufal.mff.cuni.cz