LREC2008 – W11 - ELRA Workshop on

**Evaluation Looking into the Future of Evaluation:
when automatic metrics meet task-based and
performance-based approaches**

27 May 2008 – Karam 5

WORKSHOP PROGRAMME

09:00   Welcome and Introduction

09:15   *Technology Advancement has Required Evaluations to Change Data and
Tasks -- and now Metrics*
Mark Przybocki; NIST, USA

10:00   *Explicit and Implicit Requirements of Technology Evaluations: Implications
for Test Data Creation*
Lauren Friedman, Stephanie Strassel, Meghan Lammie Glenn; Linguistic Data
Consortium, USA

10:35   Coffee break

11:00   *Automated MT Evaluation for Error Analysis: Automatic Discovery of
Potential Translation Errors for Multiword Expressions*
Bogdan Babych, Anthony Hartley; Centre for Translation Studies, University
of Leeds, UK

11:35   *Discussion*

12:15   *Reference-based vs. Task-based Evaluation of Human Language Technology*
Andrei Popescu-Belis; IDIAP Research Institute, Switzerland

12:50   *FEIRI: Extending ISLE's FEMTI for the Evaluation of a Specialized
Application in Information Retrieval*
Keith J. Miller; The MITRE Corporation, USA

13:25   Lunch

14:30   *Discussion*

15:15   *Evaluating a Natural Language Processing Approach in Arabic Information Retrieval*
Nasredine Semmar, Laib Meriama, Christian Fluhr; CEA, LIST, Laboratoire d'ingénierie de la Connaissance Multimédia Multilingue, France, NewPhenix, France

15:50   Coffee break

16:20   *A Review of the Benefits and Issues of Speaker Verification Evaluation Campaigns*
Asmaa El Hannani, Jean Hennebert; Department of Computer Science, University of Sheffield, UK, HESSO, Business Information Systems, Switzerland and University of Fribourg, Switzerland

16:55   *Field Testing of an Interactive Question-Answering Character*
Ron Artstein, Sudeep Gandhe, Anton Leuski and David Traum; Institute for Creative Technologies, University of Southern California, USA

17:30   Discussion and conclusions

18:15   Close