INVITED SPEAKER:

Hermann Ney, RWTH-Aachen

PROGRAM COMMITTEE:

Kevin Knight, USC/ISI, Organizer
Franz Josef Och, RWTH-Aachen, Organizer
Jessie Pinkham, Microsoft Research, Organizer
Deborah Coughlin, Microsoft Research, Co-organizer

Srinivas Bangalore, AT&T Research
Ralf Brown, CMU
Francisco Casacuberta, Polytechnic Univ. of Valencia
Eugene Charniak, Brown University
Robert Frederking, CMU
Ulf Hermjakob, USC/ISI
Bob Moore, Microsoft Research
Masaaki Nagata, NTT
Joseph Pentheroudakis, Microsoft Research
Norbert Reithinger, DFKI
Philip Resnik, Univ. of Maryland
Steve Richardson, Microsoft Research
Eiichiro Sumita, ATR
Koichi Takeda, IBM TR
Enrique Vidal, Polytechnic Univ. of Valencia
Stephan Vogel, Univ. of Kaiserslautern
Hideo Watanabe, IBM TRL

ADDITIONAL REVIEWERS:

Michele Banko, Microsoft Research
Marisa Jimenez, Microsoft Research
Eric Ringger, Microsoft Research
Hisami Suzuki, Microsoft Research

INTRODUCTION

This volume contains the papers accepted for presentation at the 2001 Workshop on Data Driven Machine Translation, which is part of the 39th Annual Meeting of the Association for Computational Linguistics held on July 7, 2001 in Toulouse, France.

With the increased availability of online corpora, data-driven approaches have become central to the NL community. A variety of data-driven approaches have been used to help build Machine Translation systems –example-based, statistical MT, and other machine learning approaches – and there are all sorts of possibilities for hybrid systems. We have endeavored to bring together proponents of as many techniques as possible to engage in a discussion of which combinations will yield maximal success in translation. We have chosen to center the workshop on Data Driven MT, by which we mean all approaches which develop algorithms and programs to exploit data in the development of MT, primarily the use of large bilingual corpora created by human translators, and serving as a source of training data for MT systems.

To further this theme, the conference program also includes an invited talk on the subject. We would like to thank our invited speaker, Hermann Ney, for his participation.

Organizing a workshop like this is a bit of an effort –especially given what seemed to be very tight deadlines. Luckily, we had exceptionally cooperative program committee members and reviewers. From an initial slate of 26 submissions, they selected the papers in this proceedings in a blind refereeing process. We deeply and sincerely thank all who reviewed: we would have run aground without their efforts.

Additionally, we would like to thank Deborah Coughlin for joining us in organizing the workshop. She was able to implement the reviewing process, organize the reception of the papers and master all the nitty-gritty without which the workshop and the proceedings could not happen.

We most gratefully acknowledge the support of our sponsor, Microsoft Research, for generous support of this activity.

Jessie Pinkham
Franz Josef Och
Kevin Knight
May 2001