

NAACL-HLT 2007

**Doctoral Consortium**  
**Proceedings of the Workshop**

22 April 2007  
University of Rochester  
Rochester, New York, USA

Production and Manufacturing by  
*Omnipress Inc.*  
*Post Office Box 7214*  
*Madison, WI 53707-7214*



©2007 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
75 Paterson Street, Suite 9  
New Brunswick, NJ 08901  
USA  
Tel: +1-732-342-9100  
Fax: +1-732-342-9339  
[acl@aclweb.org](mailto:acl@aclweb.org)

## **Introduction**

These are the proceedings of the NAACL-HLT 2007 Doctoral Consortium. Ten participants were selected from a total of eighteen applicants based on submission quality, advisor recommendation, and expected completion date for the PhD degree. The goals of this event are to provide these senior Ph.D. students with the opportunity to discuss and explore their research and career objectives with a panel of established researchers in the fields of natural language processing, speech technology, and information retrieval, and to develop the skills necessary to effectively communicate their research in preparation for future job talks.



**Organizers:**

Jackson Liscombe, Columbia University  
Phillip Michalak, University of Rochester

**Faculty Advisor:**

Julia Hirschberg, Columbia University

**Panelists:**

James Allen  
Chris Brew  
Ciprian Chelba  
Mona Diab  
Graeme Hirst  
Ed Hovy  
Kevin Knight  
Roland Kuhn  
Gina Levow  
Mitch Marcus  
Bob Moore  
Ani Nenkova  
Mari Ostendorf  
Michael Riley  
Brian Roark  
Stephen Robertson  
Candace Sidner  
Mark Steedman  
Ellen Voorhees  
Bonnie Webber  
Ralph Weischedel



## Table of Contents

|  |    |
|--|----|
| <i>Query Expansion Using Domain Information in Compounds</i>                                 |    |
| Karin Friberg .....  | 1  |
| <i>Learning Structured Classifiers for Statistical Dependency Parsing</i>                    |    |
| Qin Iris Wang .....  | 5  |
| <i>Creating a Knowledge Base from a Collaboratively Generated Encyclopedia</i>               |    |
| Simone Paolo Ponzetto .....  | 9  |
| <i>Knowledge-Based Labeling of Semantic Relationships in English</i>                         |    |
| Alicia Tribble .....   | 13 |
| <i>Analysis of Summarization Evaluation Experiments</i>                                      |    |
| Marie-Josée Goulet .....   | 17 |
| <i>Exploiting Event Semantics to Parse the Rhetorical Structure of Natural Language Text</i> |    |
| Rajen Subba .....  | 21 |
| <i>Dynamic Use of Ontologies in Dialogue Systems</i>   |    |
| Joana Paulo Pardal .....   | 25 |
| <i>Semantic Frames in Romanian Natural Language Processing Systems</i>                       |    |
| Diana Marie Trandabăț .....  | 29 |
| <i>Combining Evidence for Improved Speech Retrieval</i>                                      |    |
| J. Scott Olsson .....  | 33 |
| <i>Unsupervised Natural Language Processing Using Graph Models</i>                           |    |
| Chris Biemann .....  | 37 |





## Doctoral Consortium Program

**Sunday, April 22, 2007**

- 8:30–9:00 Breakfast
- 9:00–9:40 *Query Expansion Using Domain Information in Compounds*  
Karin Friberg
- 9:40–10:20 *Learning Structured Classifiers for Statistical Dependency Parsing*  
Qin Iris Wang
- 10:20–11:00 *Creating a Knowledge Base from a Collaboratively Generated Encyclopedia*  
Simone Paolo Ponzetto
- 11:00–11:40 *Knowledge-Based Labeling of Semantic Relationships in English*  
Alicia Tribble
- 11:40–12:20 *Analysis of Summarization Evaluation Experiments*  
Marie-Josée Goulet
- 12:20–13:20 Lunch
- 13:20–14:00 *Exploiting Event Semantics to Parse the Rhetorical Structure of Natural Language Text*  
Rajen Subba
- 14:00–14:40 *Dynamic Use of Ontologies in Dialogue Systems*  
Joana Paulo Pardal
- 14:40–15:20 *Semantic Frames in Romanian Natural Language Processing Systems*  
Diana Marie Trandabăţ
- 15:20–16:00 *Combining Evidence for Improved Speech Retrieval*  
J. Scott Olsson
- 16:00–16:40 *Unsupervised Natural Language Processing Using Graph Models*  
Chris Biemann
- 16:40–17:00 Coffee Break
- 17:00–18:00 Panel Discussion

