Workshop on Biomedical Natural Language Processing 2011

(BioNLP 2011)

Held at ACL HLT 2011

Portland, Oregon, USA
23-24 June 2011

ISBN: 978-1-61839-233-6

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2011) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2011)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Not all links are equal: Exploiting Dependency Types for the Extraction of Protein-Protein Interactions from Text
Philippe Thomas, Stefan Pietschmann, Illés Solt, Domonkos Tikk and Ulf Leser 1

Unsupervised Entailment Detection between Dependency Graph Fragments
Marek Rei and Ted Briscoe 10

Learning Phenotype Mapping for Integrating Large Genetic Data
Chun-Nan Hsu, Cheng-Ju Kuo, Congxing Cai, Sarah Pendergrass, Marylyn Ritchie and Jose Luis Ambite 19

EVEX: A PubMed-Scale Resource for Homology-Based Generalization of Text Mining Predictions
Sofie Van Landeghem, Filip Ginter, Yves Van de Peer and Tapio Salakoski 28

Fast and simple semantic class assignment for biomedical text
K. Bretonnel Cohen, Thomas Christiansen, William Baumgartner Jr., Karin Verspoor and Lawrence Hunter 38

The Role of Information Extraction in the Design of a Document Triage Application for Biocuration
Sandeep Pokkunuri, Cartic Ramakrishnan, Ellen Riloff, Eduard Hovy and Gully Burns 46

Medical Entity Recognition: A Comparaison of Semantic and Statistical Methods
Asma Ben Abacha and Pierre Zweigenbaum 56

Automatic Acquisition of Huge Training Data for Bio-Medical Named Entity Recognition
Yu Usami, Han-Cheol Cho, Naoaki Okazaki and Jun'ichi Tsujii 65

Building frame-based corpus on the basis of ontological domain knowledge
He Tan, Rajaram Kaliyaperumal and Nirupama Benis 74

Building a Coreference-Annotated Corpus from the Domain of Biochemistry
Riza Theresa Batista-Navarro and Sophia Ananiadou 83

Towards Morphologically Annotated Corpus of Hospital Discharge Reports in Polish
Malgorzata Marciniak and Agnieszka Mykowiecka 92

In Search of Protein Locations
Catherine Blake and Wu Zheng 101

Automatic extraction of data deposition statements: where do the research results go?
Aurelie Neveol, W. John Wilbur and Zhiyong Lu 103

From Pathways to Biomolecular Events: Opportunities and Challenges
Tomoko Ohta, Sampo Pyysalo and Jun'ichi Tsujii 105

Towards Exhaustive Event Extraction for Protein Modifications
Sampo Pyysalo, Tomoko Ohta, Makoto Miwa and Jun'ichi Tsujii 114

A Study on Dependency Tree Kernels for Automatic Extraction of Protein-Protein Interaction
Faisal Md. Chowdhury, Alberto Lavelli and Alessandro Moschitti 124

Hypothesis and Evidence Extraction from Full-Text Scientific Journal Articles
Elizabeth White, K. Bretonnel Cohen and Larry Hunter 134

SimSem: Fast Approximate String Matching in Relation to Semantic Category Disambiguation
Pontus Stenetorp, Sampo Pyysalo and Jun'ichi Tsujii 136

Building Timelines from Narrative Clinical Records: Initial Results Based-on Deep Natural Language Understanding
Hyuckchul Jung, James Allen, Nate Blaylock, William de Beaumont, Lucian Galescu and Mary Swift 146

Text Mining Techniques for Leveraging Positively Labeled Data
Lana Yeganova, Donald C. Comeau, Won Kim and W. John Wilbur 155

Parsing Natural Language Queries for Life Science Knowledge
Tadayoshi Hara, Yuka Tateisi, Jin-Dong Kim and Yusuke Miyao 164

Unlocking Medical Ontologies for Non-Ontology Experts
Shao Fen Liang, Donia Scott, Robert Stevens and Alan Rector 174

Self-training and co-training in biomedical word sense disambiguation
Antonio Jimeno Yepes and Alan Aronson 182

Medstract - The Next Generation
Marc Verhagen and James Pustejovsky 184

ThaiHerbMiner: A Thai Herbal Medicine Mining and Visualizing Tool
Choochart Haruechaiyasak, Jaruwat Pailai, Wasna Viratyosin and Rachada Kongkachandra . . 186