

# **2nd Louhi Workshop on Text and Data Mining of Health Documents 2010**

**(Louhi-10)**

**Held at NAACL HLT 2010**

**Los Angeles, California, USA  
5 June 2010**

**ISBN: 978-1-61738-680-0**

**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© (2010) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2010)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>MedEval- A Swedish Medical Test Collection with Doctors and Patients User Groups</i> Karin Friberg Heppin .....	1
<i>Extracting Information for Generating A Diabetes Report Card from Free Text in Physicians Notes</i> Ramanjot Singh Bhatia, Amber Graystone, Ross A Davies, Susan McClinton, Jason Morin and Richard F Davies .....	8
<i>Negation Detection in Swedish Clinical Text</i> Maria Skeppstedt .....	15
<i>Using Domain Knowledge about Medications to Correct Recognition Errors in Medical Report Cre- ation</i> Stephanie Schreitter, Alexandra Klein, Johannes Matiasek and Harald Trost .....	22
<i>Assessment of Utility in Web Mining for the Domain of Public Health</i> Peter von Etter, Silja Huttunen, Arto Vihavainen, Matti Vuorinen and Roman Yangarber .....	29
<i>Reliability and Type of Consumer Health Documents on the World Wide Web: an Annotation Study</i> Melanie Martin .....	38
<i>Automated Identification of Synonyms in Biomedical Acronym Sense Inventories</i> Genevieve B. Melton, SungRim Moon, Bridget McInnes and Serguei Pakhomov .....	46
<i>Characteristics and Analysis of Finnish and Swedish Clinical Intensive Care Nursing Narratives</i> Helen Allvin, Elin Carlsson, Hercules Dalianis, Riitta Danielsson-Ojala, Vidas Daudaravicius, Martin Hassel, Dimitrios Kokkinakis, Heljä Lundgren-Laine, Gunnar Nilsson, Øystein Nytrø, Sanna Salanterä, Maria Skeppstedt, Hanna Suominen and Sumithra Velupillai .....	53
<i>Extracting Medication Information from Discharge Summaries</i> Scott Halgrim, Fei Xia, Imre Solti, Eithon Cadag and Özlem Uzuner .....	61
<i>Linking SweFN++ with Medical Resources, towards a MedFrameNet for Swedish</i> Dimitrios Kokkinakis and Maria Toporowska Gronostaj .....	68
<i>Measuring Risk and Information Preservation: Toward New Metrics for De-identification of Clinical Texts</i> Lynette Hirschman and John Aberdeen .....	72
<i>A Comparison of Several Key Information Visualization Systems for Secondary Use of Electronic Health Record Content</i> Francisco Roque, Laura Slaughter and Aleksandr Tkatchenko .....	76
<i>Machine learning and features selection for semi-automatic ICD-9-CM encoding</i> Julia Medori and Cédric Fairon .....	84
<i>Extracting Formulaic and Free Text Clinical Research Articles Metadata using Conditional Random Fields</i> Sein Lin, Jun-Ping Ng, Shreyasee Pradhan, Jatin Shah, Ricardo Pietrobon and Min-Yen Kan ..	90