

6th International Workshop on Health Text Mining and Information Analysis (LOUHI 2015)

Held at EMNLP 2015

Lisbon, Portugal
17 September 2015

ISBN: 978-1-5108-1191-1

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© (2015) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2015)

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

<i>In-depth annotation for patient level liver cancer staging</i> Wen-wai Yim, Sharon Kwan and Meliha Yetisgen	1
<i>Predicting Continued Participation in Online Health Forums</i> Farig Sadeque, Thamar Solorio, Ted Pedersen, Prasha Shrestha and Steven Bethard	12
<i>Redundancy in French Electronic Health Records: A preliminary study</i> Eva D'hondt, Xavier Tannier and Aurélie Névéol	21
<i>Is it possible to recover personal health information from an automatically de-identified corpus of French EHRs?</i> Cyril Grouin, Nicolas Griffon and Aurélie Névéol	31
<i>An Analysis of Biomedical Tokenization: Problems and Strategies</i> Noa P. Cruz Diaz and Manuel Maña López	40
<i>Annotation of Clinically Important Follow-up Recommendations in Radiology Reports</i> Meliha Yetisgen, Prescott Klassen, Lucas McCarthy, Elena Pellicer, Tom Payne and Martin Gunn 50	
<i>On the Impact of Twitter-based Health Campaigns: A Cross-Country Analysis of Movember</i> Nugroho Dwi Prasetyo, Claudia Hauff, Dong Nguyen, Tijs van den Broek and Djoerd Hiemstra	55
<i>Exploring Word Embedding for Drug Name Recognition</i> Isabel Segura-Bedmar, Víctor Suárez-Paniagua and Paloma Martínez	64
<i>Creating a rule based system for text mining of Norwegian breast cancer pathology reports</i> Rebecka Weegar and Hercules Dalianis	73
<i>Parser Adaptation to the Biomedical Domain without Re-Training</i> Jeff Mitchell and Mark Steedman	79
<i>Expanding a dictionary of marker words for uncertainty and negation using distributional semantics</i> Alyaa Alfalahi, Maria Skeppstedt, Rickard Ahlbom, Roza Baskalayci, Aron Henriksson, Lars Asker, Carita Paradis and Andreas Kerren	90
<i>Held-out versus Gold Standard: Comparison of Evaluation Strategies for Distantly Supervised Relation Extraction from Medline abstracts</i> Roland Roller and Mark Stevenson	97
<i>Checking a structured pathology report for completeness of content using terminological knowledge</i> Sebastian Busse	103
<i>Effectively Crowdsourcing Radiology Report Annotations</i> Anne Cocos, Aaron Masino, Ting Qian, Ellie Pavlick and Chris Callison-Burch	109
<i>Identifying Key Concepts from EHR Notes Using Domain Adaptation</i> Jiaping Zheng and Hong Yu	115
<i>Information Extraction from Biomedical Texts: Learning Models with Limited Supervision</i> Marie-Francine Moens	120

<i>Adverse Drug Event classification of health records using dictionary based pre-processing and machine learning</i>	
Stefanie Friedrich and Hercules Dalianis	121
<i>NLP–Based Readability Assessment of Health–Related Texts: a Case Study on Italian Informed Consent Forms</i>	
Giulia Venturi, Tommaso Bellandi, Felice Dell’Orletta and Simonetta Montemagni	131
<i>Mining and Ranking Biomedical Synonym Candidates from Wikipedia</i>	
Abhyuday Jagannatha, Jinying Chen and Hong Yu	142
<i>Representing Clinical Notes for Adverse Drug Event Detection</i>	
Aron Henriksson	152