

11th International Conference on Recent Advances in Natural Language Processing (RANLP 2017)

Varna, Bulgaria
2 – 8 September 2017

ISBN: 978-1-5108-5522-9

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© (2017) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2018)

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: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

PARAMETER TRANSFER ACROSS DOMAINS FOR WORD SENSE DISAMBIGUATION	1
<i>Sallam Abualhaija ; Nina Tahmasebi ; Diane Forin ; Karl-Heinz Zimmermann</i>	
WHAT SENTENCE ARE YOU REFERRING TO AND WHY? IDENTIFYING CITED SENTENCES IN SCIENTIFIC LITERATURE	9
<i>Ahmed Abura'Ed ; Luis Chiruzzo ; Horacio Saggion</i>	
A COMPARISON OF FEATURE-BASED AND NEURAL SCANSION OF POETRY	18
<i>Manex Agirrezabal ; Inaki Alegria ; Mans Hulden</i>	
PERSIAN-SPANISH LOW-RESOURCE STATISTICAL MACHINE TRANSLATION THROUGH ENGLISH AS PIVOT LANGUAGE	24
<i>Benyamin Ahmadnia ; Javier Serrano ; Gholamreza Haffari</i>	
SIMPLE OPEN STANCE CLASSIFICATION FOR RUMOUR ANALYSIS	31
<i>Ahmet Aker ; Leon Derczynski ; Kalina Bontcheva</i>	
AN EXTENSIBLE MULTILINGUAL OPEN SOURCE LEMMATIZER	40
<i>Ahmet Aker ; Johann Petrak ; Firas Sabbah</i>	
UNIVERSAL DEPENDENCIES FOR ARABIC TWEETS	46
<i>Fahad Albogamy ; Allan Ramsay</i>	
TRANSLATING DIALECTAL ARABIC AS LOW RESOURCE LANGUAGE USING WORD EMBEDDING	52
<i>Ebtessam H Almansor ; Ahmed Al-Ani</i>	
USING ENGLISH DICTIONARIES TO GENERATE COMMONSENSE KNOWLEDGE IN NATURAL LANGUAGE	58
<i>Ali Almiman ; Allan Ramsay</i>	
A HYBRID SYSTEM TO APPLY NATURAL LANGUAGE INFERENCE OVER DEPENDENCY TREES	64
<i>Ali Almiman ; Allan Ramsay</i>	
ENSEMBLES OF CLASSIFIERS FOR CLEANING WEB PARALLEL CORPORA AND TRANSLATION MEMORIES	71
<i>Eduard Barbu</i>	
EXPLOITING AND EVALUATING A SUPERVISED, MULTILINGUAGE KEYPHRASE EXTRACTION PIPELINE FOR UNDER-RESOURCED LANGUAGES	78
<i>Marco Basaldella ; Muhammad Helmy ; Elisa Antolli ; Mihai Horia Popescu ; Giuseppe Serra ; Carlo Tasso</i>	
MULTI-LINGUAL PHRASE-BASED STATISTICAL MACHINE TRANSLATION FOR ARABIC-ENGLISH	86
<i>Ahmed Bastawisy ; Mohamed Elmahdy</i>	
INTER-ANNOTATOR AGREEMENT IN SENTIMENT ANALYSIS: MACHINE LEARNING PERSPECTIVE	97
<i>Victoria Bobicev ; Marina Sokolova</i>	
FAST AND ACCURATE DECISION TREES FOR NATURAL LANGUAGE PROCESSING TASKS	103
<i>Tiberiu Boros ; Stefan Daniel Dumitrescu ; Sonia Pipa</i>	
AN EVOLUTIONARY ALGORITHM FOR AUTOMATIC SUMMARIZATION	111
<i>Aurelien Bossard ; Christophe Rodrigues</i>	
BUILDING CHATBOTS FROM FORUM DATA: MODEL SELECTION USING QUESTION ANSWERING METRICS	121
<i>Martin Boyanov ; Preslav Nakov ; Alessandro Moschitti ; Giovanni Da San Martino ; Ivan Koychev</i>	
MINING ASSOCIATION RULES FROM CLINICAL NARRATIVES	130
<i>Svetla Boytcheva ; Ivelina Nikolova ; Galia Angelova</i>	
SENTENCE-LEVEL MULTILINGUAL MULTI-MODAL EMBEDDING FOR NATURAL LANGUAGE PROCESSING	139
<i>Iacer Calixto ; Qun Liu</i>	
ROLE-BASED MODEL FOR NAMED ENTITY RECOGNITION	149
<i>Pablo Calleja ; Raul Garcia Castro ; Guadalupe Aguado-De-Cea ; Asuncion Gomez-Perez</i>	
UNDERSPECIFICATION IN NATURAL LANGUAGE UNDERSTANDING FOR DIALOG AUTOMATION	164
<i>John Chen ; Srinivas Bangalore</i>	
IDENTIFICATION AND CLASSIFICATION OF THE MOST IMPORTANT MOMENTS IN STUDENTS COLLABORATIVE CHATS	171
<i>Costin Chiru ; Remus Decea</i>	

ANNOTATION OF ENTITIES AND RELATIONS IN SPANISH RADIOLOGY REPORTS	177
<i>Viviana Cotik ; Dario Filippo ; Roland Roller ; Hans Uszkoreit ; Feiyu Xu</i>	
TOWARDS REPLICABILITY IN PARSING	185
<i>Daniel Dakota ; Sandra Kubler</i>	
ON THE STYLISTIC EVOLUTION FROM COMMUNISM TO DEMOCRACY: SOLOMON	
MARCUS STUDY CASE	201
<i>Anca Dinu ; Liviu P. Dinu ; Bogdan Dumitru</i>	
BUILDING TIMELINES OF SOCCER MATCHES FROM TWITTER	208
<i>Amosse Edouard ; Elena Cabrio ; Sara Tonelli ; Nhan Le-Than</i>	
YOU'LL NEVER TWEET ALONE: BUILDING SPORTS MATCH TIMELINES FROM	
MICROBLOG POSTS	214
<i>Amosse Edouard ; Elena Cabrio ; Sara Tonelli ; Nhan Le-Thanh</i>	
GRAPH-BASED EVENT EXTRACTION FROM TWITTER	222
<i>Amosse Edouard ; Elena Cabrio ; Sara Tonelli ; Nhan Le-Thanh</i>	
OPINION MINING IN SOCIAL NETWORKS VERSUS ELECTORAL POLLS	231
<i>Javi Fernandez ; Fernando Llopis ; Yoan Gutierrez ; Patricio Martinez-Barco ; Alvaro Diez</i>	
CORPUS CREATION AND INITIAL SMT EXPERIMENTS BETWEEN SPANISH AND SHIPIBO-	
KONIBO	238
<i>Ana Paula Galarreta ; Andres Melgar ; Arturo Oncevay</i>	
RUSSIAN-TATAR SOCIO-POLITICAL THESAURUS: METHODOLOGY, CHALLENGES, THE	
STATUS OF THE PROJECT	245
<i>Alfiya Galieva ; Olga Nevzorova ; Dilyara Yakubova</i>	
ON A CHAT BOT FINDING ANSWERS WITH OPTIMAL RHETORIC REPRESENTATION	253
<i>Boris Galitsky ; Dmitry Ilvovsky</i>	
DETECTING ONLINE HATE SPEECH USING CONTEXT AWARE MODELS	260
<i>Lei Gao ; Ruihong Huang</i>	
A CONTEXT-AWARE APPROACH FOR DETECTING WORTH-CHECKING CLAIMS IN	
POLITICAL DEBATES	267
<i>Pepa Gencheva ; Preslav Nakov ; Lluis Marquez ; Alberto Barron-Cedeno ; Ivan Koychev</i>	
HASHTAG PROCESSING FOR ENHANCED CLUSTERING OF TWEETS	277
<i>Dagmar Gromann ; Thierry Declerck</i>	
NATURAL LANGUAGE PROCESSING TECHNOLOGIES FOR DOCUMENT PROFILING	284
<i>Antonio Guillen ; Yoan Gutierrez ; Rafael Munoz</i>	
MAPPSENT: A TEXTUAL MAPPING APPROACH FOR QUESTION-TO-QUESTION	
SIMILARITY	291
<i>Amir Hazem ; Basma El Amel Boussaha ; Nicolas Hernandez</i>	
THE IMPACT OF FIGURATIVE LANGUAGE ON SENTIMENT ANALYSIS	301
<i>Tomas Hercig ; Ladislav Lenc</i>	
ARGUMENT LABELING OF EXPLICIT DISCOURSE RELATIONS USING LSTM NEURAL	
NETWORKS	309
<i>Sohail Hooda ; Leila Kosseim</i>	
NON-DETERMINISTIC SEGMENTATION FOR CHINESE LATTICE PARSING	316
<i>Hai Hu ; Daniel Dakota ; Sandra Kubler</i>	
GOOD NEWS VS. BAD NEWS: WHAT ARE THEY TALKING ABOUT?	325
<i>Olga Kanishcheva ; Victoria Bobicev</i>	
WE BUILT A FAKE NEWS / CLICK BAIT FILTER: WHAT HAPPENED NEXT WILL BLOW	
YOUR MIND!	334
<i>Georgi Karadzhev ; Pepa Gencheva ; Preslav Nakov ; Ivan Koychev</i>	
FULLY AUTOMATED FACT CHECKING USING EXTERNAL SOURCES	344
<i>Georgi Karadzhev ; Preslav Nakov ; Lluis Marquez ; Alberto Barron-Cedeno ; Ivan Koychev</i>	
MAKING TRAVEL SMARTER: EXTRACTING TRAVEL INFORMATION FROM EMAIL	
ITINERARIES USING NAMED ENTITY RECOGNITION	354
<i>Divyansh Kaushik ; Shashank Gupta ; Chakradhar Raju ; Reuben Dias ; Sanjib Ghosh</i>	
GRAPH-BASED APPROACH TO RECOGNIZING CST RELATIONS IN POLISH TEXTS	363
<i>Pawel Kedzia ; Maciej Piasecki ; Arkadiusz Janz</i>	
DOMAIN CONTROL FOR NEURAL MACHINE TRANSLATION	372
<i>Catherine Kobus ; Josep Crego ; Jean Senellart</i>	
CURRICULUM LEARNING AND MINIBATCH BUCKETING IN NEURAL MACHINE	
TRANSLATION	379
<i>Tom Kocmi ; Ondrej Bojar</i>	
IMPROVED RECOGNITION AND NORMALISATION OF POLISH TEMPORAL EXPRESSIONS	387
<i>Jan Koco ; Michal Marciczuk</i>	

JOINT UNSUPERVISED LEARNING OF SEMANTIC REPRESENTATION OF WORDS AND ROLES IN DEPENDENCY TREES	394
<i>Michal Konkol</i>	
CZECH DATASET FOR SEMANTIC SIMILARITY AND RELATEDNESS	401
<i>Miloslav Konopik ; Ondrej Prazak ; David Steinberger</i>	
IMPROVING DISCOURSE RELATION PROJECTION TO BUILD DISCOURSE ANNOTATED CORPORA	407
<i>Majid Laali ; Leila Kosseim</i>	
EXTRACTING SEMANTIC RELATIONS VIA THE COMBINATION OF INFERENCES, SCHEMAS AND COOCCURRENCES	417
<i>Mathieu Lafourcade ; Nathalie Le Brun</i>	
IF MICE WERE REPTILES, THEN REPTILES COULD BE MAMMALS OR HOW TO DETECT ERRORS IN THE JEUXDEMOTS LEXICAL NETWORK?	424
<i>Mathieu Lafourcade ; Alain Joubert ; Nathalie Le Brun</i>	
WORD EMBEDDINGS FOR MULTI-LABEL DOCUMENT CLASSIFICATION	431
<i>Ladislav Lenc ; Pavel Kral</i>	
GENDER PREDICTION FOR CHINESE SOCIAL MEDIA DATA	438
<i>Wen Li ; Markus Dickinson</i>	
A STATISTICAL MACHINE TRANSLATION MODEL WITH FOREST-TO-TREE ALGORITHM FOR SEMANTIC PARSING	446
<i>Zhihua Liao ; Yan Xie</i>	
SUMMARIZING WORLD SPEAK : A PRELIMINARY GRAPH BASED APPROACH	452
<i>Nikhil Londhe ; Rohini Srihari</i>	
HUMAN ASSOCIATIONS HELP TO DETECT CONVENTIONALIZED MULTIWORD EXPRESSIONS	459
<i>Natalia Loukachevitch ; Anastasia Gerasimova</i>	
DETECTING HATE SPEECH IN SOCIAL MEDIA	467
<i>Shervin Malmasi ; Marcos Zampieri</i>	
INFOREX A COLLABORATIVE SYSTEM FOR TEXT CORPORA ANNOTATION AND ANALYSIS	473
<i>Michal Marciczuk ; Marcin Oleksy ; Jan Koco</i>	
LEMMATIZATION OF MULTI-WORD COMMON NOUN PHRASES AND NAMED ENTITIES IN POLISH	483
<i>Michal Marciczuk</i>	
LOG-LINEAR MODELS FOR UYGHUR SEGMENTATION IN SPOKEN LANGUAGE TRANSLATION	492
<i>Chenggang Mi ; Yating Yang ; Rui Dong ; Xi Zhou ; Lei Wang ; Xiao Li ; Tonghai Jiang</i>	
BOOTSTRAPPING A ROMANIAN CORPUS FOR MEDICAL NAMED ENTITY RECOGNITION	501
<i>Maria Mitrofan</i>	
A DOMAIN AND LANGUAGE INDEPENDENT NAMED ENTITY CLASSIFICATION APPROACH BASED ON PROFILES AND LOCAL INFORMATION	510
<i>Isabel Moreno ; Maria Teresa Roma-Ferri ; Paloma Moreda Pozo</i>	
SIMILARITY BASED GENRE IDENTIFICATION FOR POS TAGGING EXPERTS & DEPENDENCY PARSING	519
<i>Atreyee Mukherjee ; Sandra Kubler</i>	
RECOGNIZING REPUTATION DEFENCE STRATEGIES IN CRITICAL POLITICAL EXCHANGES	527
<i>Nona Naderi ; Graeme Hirst</i>	
CLASSIFYING FRAMES AT THE SENTENCE LEVEL IN NEWS ARTICLES	536
<i>Nona Naderi ; Graeme Hirst</i>	
ROBUST TUNING DATASETS FOR STATISTICAL MACHINE TRANSLATION	543
<i>Preslav Nakov ; Stephan Vogel</i>	
DO NOT TRUST THE TROLLS: PREDICTING CREDIBILITY IN COMMUNITY QUESTION ANSWERING FORUMS	551
<i>Preslav Nakov ; Tsvetomila Mihaylova ; Lluís Marquez ; Yashkumar Shiroya ; Ivan Koychev</i>	
BULGARIAN-ENGLISH AND ENGLISH-BULGARIAN MACHINE TRANSLATION: SYSTEM DESIGN AND EVALUATION	561
<i>Petya Osenova ; Kiril Simov</i>	
IDENTIFICATION OF CHARACTER ADJECTIVES FROM MAHABHARATA	569
<i>Apurba Paul ; Dipankar Das</i>	
LEARNING MULTIMODAL GENDER PROFILE USING NEURAL NETWORKS	577
<i>Carlos Perez Estruch ; Roberto Paredes Palacios ; Paolo Rosso</i>	

RECOGNITION OF GENUINE POLISH SUICIDE NOTES	583
<i>Maciej Piasecki ; Ksenia Mlynarczyk ; Jan Koco</i>	
CROSS-LINGUAL SRL BASED UPON UNIVERSAL DEPENDENCIES	592
<i>Ondrej Prazak ; Miloslav Konopik</i>	
USING GAZE DATA TO PREDICT MULTIWORD EXPRESSIONS	601
<i>Omid Rohanian ; Shiva Taslimipour ; Victoria Yaneva ; Le An Ha</i>	
REAL-TIME NEWS SUMMARIZATION WITH ADAPTATION TO MEDIA ATTENTION	610
<i>Andreas Ruckle ; Iryna Gurevych</i>	
MEASURING THE LIMIT OF SEMANTIC DIVERGENCE FOR ENGLISH TWEETS.	618
<i>Dwijen Rudrapal ; Amitava Das</i>	
EVALUATING THE MORPHOLOGICAL COMPOSITIONALITY OF POLARITY	625
<i>Josef Ruppenhofer ; Petra Steiner ; Michael Wiegand</i>	
INTRODUCING EVALD -- SOFTWARE APPLICATIONS FOR AUTOMATIC EVALUATION OF DISCOURSE IN CZECH	634
<i>Katerina Rysova ; Magdalena Rysova ; Jiri Mirovsky ; Michal Novak</i>	
IDIOM TYPE IDENTIFICATION WITH SMOOTHED LEXICAL FEATURES AND A MAXIMUM MARGIN CLASSIFIER	642
<i>Giancarlo Salton ; Robert Ross ; John Kelleher</i>	
A CALIBRATION METHOD FOR EVALUATION OF SENTIMENT ANALYSIS	652
<i>F. Sharmila Sathar ; Roger Evans ; Gulden Uchyigit</i>	
BUILDING MULTIWORD EXPRESSIONS BILINGUAL LEXICONS FOR DOMAIN ADAPTATION OF AN EXAMPLE-BASED MACHINE TRANSLATION SYSTEM	661
<i>Nasredine Semmar ; Mariama Laib</i>	
IDENTIFYING THE AUTHORS' NATIONAL VARIETY OF ENGLISH IN SOCIAL MEDIA TEXTS	671
<i>Vasiliki Simaki ; Panagiotis Simakis ; Carita Paradis ; Andreas Kerren</i>	
TOWARDS LEXICAL CHAINS FOR KNOWLEDGE-GRAPH-BASED WORD EMBEDDINGS	679
<i>Kiril Simov ; Svetla Boytcheva ; Petya Osenova</i>	
WORD EMBEDDINGS AS FEATURES FOR SUPERVISED COREFERENCE RESOLUTION	686
<i>Iliana Simova ; Hans Uszkoreit</i>	
CROSS-LINGUAL FLAMES DETECTION IN NEWS DISCUSSIONS	694
<i>Josef Steinberger ; Tomas Brychcin ; Tomas Hercig ; Peter Krejzl</i>	
PYRAMID-BASED SUMMARY EVALUATION USING ABSTRACT MEANING REPRESENTATION	701
<i>Josef Steinberger ; Peter Krejzl ; Tomas Brychcin</i>	
LARGE-SCALE NEWS ENTITY SENTIMENT ANALYSIS	707
<i>Ralf Steinberger ; Stefanie Hegele ; Hristo Tanev ; Leonida Della Rocca</i>	
PREDICTING THE LAW AREA AND DECISIONS OF FRENCH SUPREME COURT CASES	716
<i>Octavia-Maria Sulea ; Marcos Zampieri ; Mihaela Vela ; Josef Van Genabith</i>	
UNSUPERVISED LEARNING OF MORPHOLOGY WITH GRAPH SAMPLING	723
<i>Maciej Sumalvico</i>	
MULTI-ENTITY SENTIMENT ANALYSIS USING ENTITY-LEVEL FEATURE EXTRACTION AND WORD EMBEDDINGS APPROACH	733
<i>Colm Sweeney ; Deepak Padmanabhan</i>	
FINDING INDIVIDUAL WORD SENSE CHANGES AND THEIR DELAY IN APPEARANCE	741
<i>Nina Tahmasebi ; Thomas Risse</i>	
STREAMING TEXT ANALYTICS FOR REAL-TIME EVENT RECOGNITION	750
<i>Philippe Thomas ; Johannes Kirschnick ; Leonhard Hennig ; Renlong Ai ; Sven Schmeier ; Holmer Hensen ; Feiyu Xu ; Hans Uszkoreit</i>	
AN EYE-TRACKING STUDY OF NAMED ENTITY ANNOTATION	758
<i>Takenobu Tokunaga ; Hitoshi Nishikawa ; Tomoya Iwakura</i>	
A GRAPH-BASED TEXT SIMILARITY MEASURE THAT EMPLOYS NAMED ENTITY INFORMATION	765
<i>Leonidas Tsekouras ; Iraklis Varlamis ; George Giannakopoulos</i>	
DETECTING METAPHORICAL PHRASES IN THE POLISH LANGUAGE	772
<i>Aleksander Wawer ; Agnieszka Mykowiecka</i>	
EFFICIENT ENCODING OF PATHOLOGY REPORTS USING NATURAL LANGUAGE PROCESSING	778
<i>Rebecka Weegar ; Jan F Nygard ; Hercules Dalianis</i>	
NEURAL RERANKING FOR NAMED ENTITY RECOGNITION	784
<i>Jie Yang ; Yue Zhang ; Fei Dong</i>	

ONLINE DECEPTION DETECTION REFUELED BY REAL WORLD DATA COLLECTION	793
<i>Wenlin Yao ; Zeyu Dai ; Ruihong Huang ; James Caverlee</i>	
A WEAKLY SUPERVISED APPROACH TO TRAIN TEMPORAL RELATION CLASSIFIERS AND ACQUIRE REGULAR EVENT PAIRS SIMULTANEOUSLY	803
<i>Wenlin Yao ; Saipravalika Nettyam ; Ruihong Huang</i>	
MULTILINGUAL AND CROSS-LINGUAL COMPLEX WORD IDENTIFICATION	813
<i>Seid Muhie Yimam ; Sanja Stajner ; Martin Riedl ; Chris Biemann</i>	
AUTOMATIC GENERATION OF SITUATION MODELS FOR PLAN RECOGNITION PROBLEMS	823
<i>Kristina Yordanova</i>	
USING NLP FOR ENHANCING SECOND LANGUAGE ACQUISITION	839
<i>Leonardo Zilio ; Rodrigo Wilkens ; Cedrick Fairon</i>	
Author Index	