

2nd International Workshop on Natural Language Processing for Digital Humanities (NLP4DH)

Online
20 November 2022

ISBN: 978-1-7138-6317-5

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© (2022) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2023)

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

<i>A Stylometric Analysis of Amadís de Gaula and Sergas de Esplandián</i> Yoshifumi Kawasaki	1
<i>Computational Exploration of the Origin of Mood in Literary Texts</i> Emily Öhman and Riikka H. Rossi	8
<i>Sentiment is all you need to win US Presidential elections</i> Sovesh Mohapatra and Somesh Mohapatra	15
<i>Interactive Analysis and Visualisation of Annotated Collocations in Spanish (AVAnCES)</i> Simon Gonzalez	21
<i>Fractality of sentiment arcs for literary quality assessment: The case of Nobel laureates</i> Yuri Bizzoni, Kristoffer Laigaard Nielbo and Mads Rosendahl Thomsen	31
<i>Style Classification of Rabbinic Literature for Detection of Lost Midrash Tanhuma Material</i> Solomon Tannor, Nachum Dershowitz and Moshe Lavee	42
<i>Use the Metadata, Luke! – An Experimental Joint Metadata Search and N-gram Trend Viewer for Personal Web Archives</i> Balázs Indig, Zsófia Sárközi-Lindner and Mihály Nagy	47
<i>MALM: Mixing Augmented Language Modeling for Zero-Shot Machine Translation</i> Kshitij Gupta	53
<i>ParsSimpleQA: The Persian Simple Question Answering Dataset and System over Knowledge Graph</i> Hamed Babaei Giglou, Niloufar Beyranvand, Reza Moradi, Amir Mohammad Salehoof and Saeed Bibak	59
<i>Enhancing Digital History – Event discovery via Topic Modeling and Change Detection</i> King Ip Lin and Sabrina Peng	69
<i>A Parallel Corpus and Dictionary for Amis-Mandarin Translation</i> Francis Zheng, Edison Marrese-Taylor and Yutaka Matsuko	79
<i>Machines in the media: semantic change in the lexicon of mechanization in 19th-century British newspapers</i> Nilo Pedrazzini and Barbara McGillivray	85
<i>Optimizing the weighted sequence alignment algorithm for large-scale text similarity computation</i> Maciej Janicki	96
<i>Domain-specific Evaluation of Word Embeddings for Philosophical Text using Direct Intrinsic Evaluation</i> Goya van Boven and Jelke Bloem	101
<i>Towards Bootstrapping a Chatbot on Industrial Heritage through Term and Relation Extraction</i> Mihael Arcan, Rory O’Halloran, Cécile Robin and Paul Buitelaar	108
<i>Non-Parametric Word Sense Disambiguation for Historical Languages</i> Enrique Manjavacas Arevalo and Lauren Fonteyn	123

<i>Introducing a Large Corpus of Tokenized Classical Chinese Poems of Tang and Song Dynasties</i> Chao-Lin Liu, Ti-Yong Zheng, Kuan-Chun Chen and Meng-Han Chung	135
<i>Creative Text-to-Image Generation: Suggestions for a Benchmark</i> Irene Russo	145
<i>The predictability of literary translation</i> Andrew Piper and Matt Erlin	155
<i>Emotion Conditioned Creative Dialog Generation</i> Khalid Alnajjar and Mika Hämmäläinen	161
<i>Integration of Named Entity Recognition and Sentence Segmentation on Ancient Chinese based on Siku-BERT</i> Sijia Ge	167
<i>(Re-)Digitizing Ngô Sĩú-lé’s Mandarin – Taiwanese Dictionary</i> Pierre Magistry and Afala Phaxay	174