4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL-2020)

Held online due to COVID-19

Barcelona, Spain
12 December 2020

ISBN: 978-1-7138-2821-1

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© (2020) by the Association for Computational Linguistics
All rights reserved.
Copyright for individual papers remains with the authors and are licensed under a Creative Commons 4.0 license. (https://creativecommons.org/licenses/by-nd/4.0/)

Printed with permission by Curran Associates, Inc. (2021)

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

History to Myths: Social Network Analysis for Comparison of Stories over Time
Clément Besnier 1

Automatic Topological Field Identification in (Historical) German Texts
Katrin Ortmann 10

Exhaustive Entity Recognition for Coptic: Challenges and Solutions
Amir Zeldes, Lance Martin and Sichang Tu 19

A Survey on Approaches to Computational Humor Generation
Miriam Amin and Manuel Burghardt 29

Neural Machine Translation of Artwork Titles Using Iconclass Codes
Nikolay Banar, Walter Daelemans and Mike Kestemont 42

A Two-Step Approach for Automatic OCR Post-Correction
Robin Schaefer and Clemens Neudecker 52

"Shakespeare in the Vectorian Age" – An evaluation of different word embeddings and NLP parameters for the detection of Shakespeare quotes
Bernhard Liebl and Manuel Burghardt 58

Vital Records: Uncover the past from historical handwritten records
Herve Dejean and Jean-Luc Meunier 69

Measuring the Effects of Bias in Training Data for Literary Classification
Sunyam Bagga and Andrew Piper 74

ERRANT: Assessing and Improving Grammatical Error Type Classification
Katerina Korre and John Pavlopoulos 85

Life still goes on: Analysing Australian WW1 Diaries through Distant Reading
Ashley Dennis-Henderson, Matthew Roughan, Lewis Mitchell and Jonathan Tuke 90

Zero-shot cross-lingual identification of direct speech using distant supervision
Murathan Kurfalı and Mats Wirén 105

Twenty-two Historical Encyclopedias Encoded in TEI: a New Resource for the Digital Humanities
Thora Hagen, Erik Ketzan, Fotis Jannidis and Andreas Witt 112

Results of a Single Blind Literary Taste Test with Short Anonymized Novel Fragments
Andreas van Cranenburgh and Corina Koolen 121

Geometric Deep Learning Models for Linking Character Names in Novels
Marek Kubis 127

Sonnet Combinatorics with OuPoCo
Thierry Poibeau, Mylène Maignant, Frédérique Mélanie-Becquet, Clément Plancq, Matthieu Raffard and Mathilde Roussel 133

Interpretation of Sentiment Analysis in Aeschylus's Greek Tragedy
Vijaya Kumari Yeruva, Mayanka ChandraShekar, Yugyung Lee, Jeff Rydberg-Cox, Virginia Blanton and Nathan A Oyler 138

Towards Olfactory Information Extraction from Text: A Case Study on Detecting Smell Experiences in Novels
Ryan Brate, Paul Groth and Marieke van Erp 147

Finding and Generating a Missing Part for Story Completion
Yusuke Mori, Hiroaki Yamane, Yusuke Mukuta and Tatsuya Harada 156

TL-Explorer: A Digital Humanities Tool for Mapping and Analyzing Translated Literature
Alex Zhai, Zheng Zhang, Amel Fraisse, Ronald Jenn, Shelley Fisher Fishkin and Pierre Zweigenbaum 167