Third Workshop on Evaluation and Comparison of NLP Systems (Eval4NLP 2022)

Online
20 November 2022

ISBN: 978-1-7138-6315-1

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© (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

A Japanese Corpus of Many Specialized Domains for Word Segmentation and Part-of-Speech Tagging
Shohei Higashiyama, Masao Ideuchi, Masao Utiyama, Yoshiaki Oida and Eiichiro Sumita 1

Assessing Resource-Performance Trade-off of Natural Language Models using Data Envelopment Analysis
Zachary Zhou, Alisha Zachariah, Devin Conathan and Jeffery Kline 11

From COMET to COMES – Can Summary Evaluation Benefit from Translation Evaluation?
Mateusz Krubiński and Pavel Pecina 21

Better Smatch = Better Parser? AMR evaluation is not so simple anymore
Juri Opitz and Anette Frank 32

GLARE: Generative Left-to-right AdversaRial Examples
Ryan Andrew Chi, Nathan Kim, Patrick Liu, Zander Lack and Ethan A Chi 44

Random Text Perturbations Work, but not Always
Zhengxiang Wang 51

A Comparative Analysis of Stance Detection Approaches and Datasets
Parush Gera and Tempestt Neal 58

Why is sentence similarity benchmark not predictive of application-oriented task performance?
Kaori Abe, Sho Yokoi, Tomoyuki Kajiwara and Kentaro Inui 70

Chat Translation Error Detection for Assisting Cross-lingual Communications
Yunmeng Li, Jun Suzuki, Makoto Morishita, Kaori Abe, Ryoko Tokuhisa, Ana Brassard and Kentaro Inui 88

Evaluating the role of non-lexical markers in GPT-2's language modeling behavior
Roberta Rocca and Alejandro de la Vega 96

Assessing Neural Referential Form Selectors on a Realistic Multilingual Dataset
Guanyi Chen, Fahime Same and Kees Van Deemter 103