

# **4th Workshop on Gender Bias in Natural Language Processing (GeBNLP 2022)**

Online  
15 July 2022

ISBN: 978-1-7138-5633-7

**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. (2022)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

|  |     |
|--|-----|
| <i>Analyzing Hate Speech Data along Racial, Gender and Intersectional Axes</i><br>Antonis Maronikolakis, Philip Baader and Hinrich Schütze .....   | 1   |
| <i>Analysis of Gender Bias in Social Perception and Judgement Using Chinese Word Embeddings</i><br>Jiali Li, Shucheng Zhu, Ying Liu and Pengyuan Liu .....   | 8   |
| <i>Don't Forget About Pronouns: Removing Gender Bias in Language Models Without Losing Factual Gender Information</i><br>Tomasz Limisiewicz and David Mareček .....  | 17  |
| <i>Uncertainty and Inclusivity in Gender Bias Annotation: An Annotation Taxonomy and Annotated Datasets of British English Text</i><br>Lucy Havens, Beatrice Alex, Benjamin Bach and Melissa Terras .....        | 30  |
| <i>Debiasing Neural Retrieval via In-batch Balancing Regularization</i><br>Yuantong Li, Xiaokai Wei, Zijian Wang, Shen Wang, Parminder Bhatia, Xiaofei Ma and Andrew Arnold .....                                | 58  |
| <i>Gender Biases and Where to Find Them: Exploring Gender Bias in Pre-Trained Transformer-based Language Models Using Movement Pruning</i><br>Przemyslaw Joniak and Akiko Aizawa .....                           | 67  |
| <i>Gendered Language in Resumes and its Implications for Algorithmic Bias in Hiring</i><br>Prasanna Parasurama and João Sedoc .....  | 74  |
| <i>The Birth of Bias: A case study on the evolution of gender bias in an English language model</i><br>Oskar Van Der Wal, Jaap Jumelet, Katrin Schulz and Willem Zuidema .....                                   | 75  |
| <i>Challenges in Measuring Bias via Open-Ended Language Generation</i><br>Afra Feyza Akyürek, Muhammed Yusuf Kocyigit, Sejin Paik and Derry Tanti Wijaya .....   | 76  |
| <i>Worst of Both Worlds: Biases Compound in Pre-trained Vision-and-Language Models</i><br>Tejas Srinivasan and Yonatan Bisk .....  | 77  |
| <i>Assessing Group-level Gender Bias in Professional Evaluations: The Case of Medical Student End-of-Shift Feedback</i><br>Emmy Liu, Michael Henry Tessler, Nicole Dubosh, Katherine Hiller and Roger Levy ..... | 86  |
| <i>On the Dynamics of Gender Learning in Speech Translation</i><br>Beatrice Savoldi, Marco Gaido, Luisa Bentivogli, Matteo Negri and Marco Turchi .....  | 94  |
| <i>Fewer Errors, but More Stereotypes? The Effect of Model Size on Gender Bias</i><br>Yarden Tal, Inbal Magar and Roy Schwartz .....   | 112 |
| <i>Unsupervised Mitigating Gender Bias by Character Components: A Case Study of Chinese Word Embedding</i><br>Xiuying Chen, Mingzhe Li, Rui Yan, Xin Gao and Xiangliang Zhang .....                              | 121 |
| <i>An Empirical Study on the Fairness of Pre-trained Word Embeddings</i><br>Emeralda Sesari, Max Hort and Federica Sarro .....   | 129 |
| <i>Mitigating Gender Stereotypes in Hindi and Marathi</i><br>Neeraja Kirtane and Tanvi Anand .....   | 145 |

|  |     |
|--|-----|
| <i>Choose Your Lenses: Flaws in Gender Bias Evaluation</i>   |     |
| Hadas Orgad and Yonatan Belinkov .....   | 151 |
| <i>A Taxonomy of Bias-Causing Ambiguities in Machine Translation</i>   |     |
| Michal Měchura .....   | 168 |
| <i>On Gender Biases in Offensive Language Classification Models</i>  |     |
| Sanjana Marcé and Adam Poliak .....  | 174 |
| <i>Gender Bias in BERT - Measuring and Analysing Biases through Sentiment Rating in a Realistic Downstream Classification Task</i> |     |
| Sophie Jentzsch and Cigdem Turan .....   | 184 |
| <i>Occupational Biases in Norwegian and Multilingual Language Models</i>   |     |
| Samia Touileb, Lilja Øvrelid and Erik Veldal .....   | 200 |
| <i>Looking for a Handsome Carpenter! Debiasing GPT-3 Job Advertisements</i>  |     |
| Conrad Borchers, Dalia Gala, Benjamin Gilbert, Eduard Oravkin, Wilfried Bounsi, Yuki M Asano and Hannah Kirk .....                 | 212 |
| <i>HeteroCorpus: A Corpus for Heteronormative Language Detection</i>   |     |
| Juan Vásquez, Gemma Bel-Enguix, Scott Andersen and Sergio-Luis Ojeda-Trueba .....  | 225 |
| <i>Evaluating Gender Bias Transfer from Film Data</i>  |     |
| Amanda Bertsch, Ashley Oh, Sanika Natu, Swetha Gangu, Alan W. Black and Emma Strubell  | 235 |
| <i>Indigenous Language Revitalization and the Dilemma of Gender Bias</i>   |     |
| Oussama Hansal, Ngoc Tan Le and Fatiha Sadat .....   | 244 |
| <i>What changed? Investigating Debiasing Methods using Causal Mediation Analysis</i>   |     |
| Sullam Jeoung and Jana Diesner .....   | 255 |
| <i>Why Knowledge Distillation Amplifies Gender Bias and How to Mitigate from the Perspective of DistilBERT</i>                     |     |
| Jaimeen Ahn, Hwaran Lee, Jinhwa Kim and Alice Oh .....   | 266 |
| <i>Incorporating Subjectivity into Gendered Ambiguous Pronoun (GAP) Resolution using Style Transfer</i>                            |     |
| Kartikey Pant and Tanvi Dadu .....   | 273 |