

Dirk Hovy

COMPUTATIONAL SOCIAL SCIENCE

ARTICLES IN SCHOLARLY JOURNALS

BOSE P., LUPO L., HABIBI M., HOVY D., SCHWARZ C.

Beyond the Stats: Realities, Perception, and Social Media Discourse on Poverty

American Economic Association Papers and Proceedings, 2024, vol.114, pp.690-694

HOVY D., MELUMAD S., INMAN J. J.

Wordify: A Tool for Discovering and Differentiating Consumer Vocabularies

Journal of Consumer Research, 2021, vol.48, no. 3, pp.394-414

PURSCHKE C., HOVY D.

Lörres, Möppes, and the Swiss. (Re)Discovering regional patterns in anonymous social media data

Journal of Linguistic Geography, 2019, vol.7, no. 2, pp.113-134

PAUN S., CARPENTER B., CHAMBERLAIN J., HOVY D., KRUSCHWITZ U., POESIO M.

Comparing Bayesian Models of Annotation

Transactions of the Association for Computational Linguistics, 2018, vol.6, pp.571-585

RESEARCH MONOGRAPHS

HOVY D.

Text Analysis in Python for Social Scientists: Prediction and Classification

Cambridge University Press, , 2022

HOVY D.

Text Analysis in Python for Social Scientists: Discovery and Exploration

Cambridge University Press, , 2020

CONTRIBUTION TO CHAPTERS, BOOKS OR RESEARCH MONOGRAPHS

BIANCHI F., FORNACIARI T., HOVY D., NOZZA D.

Gender and Age Bias in Commercial Machine Translation in *Towards Responsible Machine Translation*

Helena Moniz, Carla Parra Escartín (Eds), Springer International Publishing, pp.159-184, 2023

HOVY D., RAHIMI A., BALDWIN T., BROOKE J.

Visualizing Regional Language Variation Across Europe on Twitter in *Handbook of the Changing World*

Language Map

Stanley D Brunn, Roland Kehrein (Eds), Springer International Publishing, pp.1-24, 2019

HOVY D., QUIST P.

Computerlingvistik. Metoder til visualisering af regional variation i sociale medier in *Sociale Medier Og Sprog*
Mortensen Kristine, Stæhr Andreas (Eds), Samfunds Litteratur, chap. 5, pp.121-144, 2018

ARTICLES UPON REQUEST/INVITATION

HOVY D., PRABHUMOYE S.

Five sources of bias in natural language processing

2021, Language and Linguistics Compass

OTHER

HOVY D.

IA & società

2024, Eco, Italy

HOVY D., CIRRINCIONE A.

AI e sport business: le regole del gioco sono cambiate

2023, Capital, Italy

HOVY D., CIRRINCIONE A.

La manifattura è diventata intelligente

2023, Capital, Italy

PROCEEDINGS/PRESENTATIONS

WANG X., MA B., HU C., WEBER-GENZEL L., RÖTTGER P., KREUTER F., HOVY D., PLANK B.

“My Answer is C”: First-Token Probabilities Do Not Match Text Answers in Instruction-Tuned Language Models

Findings of the Association for Computational Linguistics ACL 2024, 2024, Bangkok, Thailand

PLAZA DEL ARCO F. M., NOZZA D., HOVY D.

Wisdom of Instruction-Tuned Language Model Crowds. Exploring Model Label Variation

Proceedings of the 3rd Workshop on Perspectivist Approaches to NLP (NLPerspectives) @ LREC-COLING 2024, May 20-25, 2024, Torino, Italy

PLAZA DEL ARCO F. M., CERCAS CURRY A., CERCAS CURRY A., ABERCROMBIE G., HOVY D.

Angry Men, Sad Women: Large Language Models Reflect Gendered Stereotypes in Emotion Attribution

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) 2024, Bangkok, Thailand

CERCAS CURRY A., ATTANASIO G., TALAT Z., HOVY D.

Classist Tools: Social Class Correlates with Performance in NLP

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) 2024, Bangkok, Thailand

SONI N., BALASUBRAMANIAN N., SCHWARTZ H. A., HOVY D.

Comparing Pre-trained Human Language Models: Is it Better with Human Context as Groups, Individual Traits, or Both?

Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis 2024, Bangkok, Thailand

PERNISI F., HOVY D., RÖTTGER P.

Compromesso! Italian Many-Shot Jailbreaks undermine the safety of Large Language Models

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop), 2024, Bangkok, Thailand

PLAZA DEL ARCO F. M., CURRY A. C., PAOLI S., CERCAS CURRY A., HOVY D.

Divine LLaMAs: Bias, Stereotypes, Stigmatization, and Emotion Representation of Religion in Large Language Models

Findings of the Association for Computational Linguistics: EMNLP 2024 2024, Miami, FL, United States of America

SINELNIK A., HOVY D.

Narratives at Conflict: Computational Analysis of News Framing in Multilingual Disinformation Campaigns

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop), 2024, Bangkok, Thailand

RÖTTGER P., HOFMANN V., PYATKIN V., HINCK M., KIRK H., SCHUETZE H., HOVY D.

Political Compass or Spinning Arrow? Towards More Meaningful Evaluations for Values and Opinions in Large Language Models

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) 2024, Bangkok, Thailand

ATTANASIO G., SAVOLDI B., FUCCI D., HOVY D.

Twists, Humps, and Pebbles: Multilingual Speech Recognition Models Exhibit Gender Performance Gaps

Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing 2024, Miami, FL, United States of America

RÖTTGER P., KIRK H., VIDGEN B., ATTANASIO G., BIANCHI F., HOVY D.

XSTest: A Test Suite for Identifying Exaggerated Safety Behaviours in Large Language Models

Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), 2024, Mexico City, Mexico

CERCAS CURRY A., ATTANASIO G., NOZZA D., HOVY D.

MilaNLP at SemEval-2023 Task 10: Ensembling Domain-Adapted and Regularized Pretrained Language Models for Robust Sexism Detection

Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023), July 13-14 2023, Toronto, Canada

ORLIKOWSKI M., RÖTTGER P., CIMIANO P., HOVY D.

The ecological fallacy in annotation: modeling human label variation goes beyond sociodemographics

Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, July 9-14 2023, Toronto, Canada

LAUSCHER A., NOZZA D., CROWLEY A., MILTERSEN E., HOVY D.

What about em? How commercial Machine Translation fails to handle (neo-)pronouns

Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, July 9-14 2023, Toronto, Canada

THEIL K., HOVY D., STUCKENSCHMIDT H.

Top-down influence? Predicting CEO personality and risk impact from speech transcripts

Proceedings of the Seventeenth International AAAI Conference on Web and Social Media, June 5-8 2023, Limassol, Cyprus

HUNG C., LAUSCHER A., HOVY D., PONZETTO S. P., GLAVAŠ G.

Can demographic factors improve text classification? Revisiting demographic adaptation in the age of

transformers

Findings of the Association for Computational Linguistics: EACL 2023, May 2-6 2023, Dubrovnik, Croatia

HADA R., EBRAHIMI FARD A., SHUGARS S., BIANCHI F., ROSSINI P., HOVY D., TROMBLE R., TINTAREV N.

Beyond Digital "Echo Chambers": The Role of Viewpoint Diversity in Political Discussion

WSDM '23: The Sixteenth ACM International Conference on Web Search and Data Mining, 27 February - 3 March 2023, Singapore

BIANCHI F., HILLS S., ROSSINI P., HOVY D., TROMBLE R., TINTAREV N.

"It's not just hate": a multi-dimensional perspective on detecting harmful speech online

The 2022 Conference on Empirical Methods in Natural Language Processing, December 7-11, 2022, Abu Dhabi, United Arab Emirates

HESSENTHALER M., STRUBELL E., HOVY D., LAUSCHER A.

Bridging fairness and environmental sustainability in natural language processing

The 2022 Conference on Empirical Methods in Natural Language Processing, December 7-11, 2022, Abu Dhabi, United Arab Emirates

BIANCHI F., CUTRONA V., HOVY D.

Twitter-demographer: a flow-based tool to enrich Twitter data

The 2022 Conference on Empirical Methods in Natural Language Processing, December 7-11, 2022, Abu Dhabi, United Arab Emirates

ATTANASIO G., NOZZA D., PASTOR E., HOVY D.

Benchmarking Post-Hoc Interpretability Approaches for Transformer-based Misogyny Detection

Proceedings of NLP Power! The First Workshop on Efficient Benchmarking in NLP, May 22-27, 2022, Dublin, Ireland

ATTANASIO G., NOZZA D., HOVY D., BARALIS E.

Entropy-based Attention Regularization Frees Unintended Bias Mitigation from Lists

Findings of the Association for Computational Linguistics: ACL 2022, May 22-27, 2022, Dublin, Ireland

FORNACIARI T., UMA A., POESIO M., HOVY D.

Hard and Soft Evaluation of NLP models with BOOtSTrap SAMpling - BooStSa

Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, May 22-27, 2022, Dublin, Ireland

BIANCHI F., NOZZA D., HOVY D.

Language Invariant Properties in Natural Language Processing

Proceedings of NLP Power! The First Workshop on Efficient Benchmarking in NLP, May 22-27, 2022, Dublin, Ireland

NOZZA D., BIANCHI F., HOVY D.

Pipelines for Social Bias Testing of Large Language Models

Proceedings of BigScience Episode #5 -- Workshop on Challenges & Perspectives in Creating Large Language Models, May 22-27, 2022, Dublin, Ireland

BIANCHI F., NOZZA D., HOVY D.

XLM-EMO: Multilingual Emotion Prediction in Social Media Text

Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, May 22-27, 2022, Dublin, Ireland

RESEARCH NOTES OR SHORT ARTICLES IN JOURNALS

BIANCHI F., CERCAS CURRY A., HOVY D.

Viewpoint: Artificial Intelligence Accidents Waiting to Happen?

2023, Journal of Artificial Intelligence Research (JAIR)
