

# Dirk Hovy

Associate Professor



**Knowledge Group:** Marketing

**Research Domains:** Computational Social Science

**Teaching Domains:**

[dirk.hovy@unibocconi.it](mailto:dirk.hovy@unibocconi.it)

---

## Articles in Scholarly Journals

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

BOSE, P., L. LUPO, M. HABIBI, D. HOVY, C., R. SCHWARZ, "Beyond the Stats: Realities, Perception, and Social Media Discourse on Poverty", American Economic Association Papers and Proceedings, 2024, vol. 114, pp. 690-694

### **Wordify: A Tool for Discovering and Differentiating Consumer Vocabularies**

HOVY, D., S. MELUMAD, J. J. INMAN, "Wordify: A Tool for Discovering and Differentiating Consumer Vocabularies", Journal of Consumer Research, 2021, vol. 48, no. 3, pp. 394-414

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

PURSCHKE, C., D. HOVY, "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

### **Comparing Bayesian Models of Annotation**

PAUN, S., B. CARPENTER, J. CHAMBERLAIN, D. HOVY, U. KRUSCHWITZ, M. POESIO, "Comparing Bayesian Models of Annotation", Transactions of the Association for Computational Linguistics, 2018, vol. 6, pp. 571-585

---

## Research Monographs

### **Text Analysis in Python for Social Scientists: Prediction and Classification**

HOVY, D. - "Text Analysis in Python for Social Scientists: Prediction and Classification" - 2022, Cambridge University Press

### **Text Analysis in Python for Social Scientists: Discovery and Exploration**

HOVY, D. - "Text Analysis in Python for Social Scientists: Discovery and Exploration" - 2020, Cambridge University Press

---

## Contribution to Chapters, Books or Research Monographs

### **Gender and Age Bias in Commercial Machine Translation**

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

### **Visualizing Regional Language Variation Across Europe on Twitter**

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

### **Computerlingvistik. Metoder til visualisering af regional variation i sociale medier**

HOVY, D., P. QUIST, "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

---

## Proceedings/Presentations

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

PLAZA DEL ARCO, F. M., A. CERCAS CURRY, A. CERCAS CURRY, G. ABERCROMBIE, D. HOVY, "Angry Men, Sad Women: Large Language Models Reflect Gendered Stereotypes in Emotion Attribution" in Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2024, Bangkok, Thailand

### **Classist Tools: Social Class Correlates with Performance in NLP**

CERCAS CURRY, A., G. ATTANASIO, Z. TALAT, D. HOVY, "Classist Tools: Social Class Correlates with Performance in NLP" in Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2024, Bangkok, Thailand

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

SONI, N., N. BALASUBRAMANIAN, H. A. SCHWARTZ, D. HOVY, "Comparing Pre-trained Human Language Models: Is it Better with Human Context as Groups, Individual Traits, or Both?" in Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, 2024, Bangkok, Thailand

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

PERNISI, F., D. HOVY, P. RÖTTGER, "Compromesso! Italian Many-Shot Jailbreaks undermine the safety of Large Language Models" in Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop), 2024, Bangkok, Thailand

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

PLAZA DEL ARCO, F. M., A. C. CURRY, S. PAOLI, A. CERCAS CURRY, D. HOVY, "Divine LLaMAs: Bias, Stereotypes, Stigmatization, and Emotion Representation of Religion in Large Language Models" in Findings of the Association for Computational Linguistics: EMNLP 2024, 2024, Miami, FL, United States of America

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

SINELNIK, A., D. HOVY, "Narratives at Conflict: Computational Analysis of News Framing in Multilingual Disinformation Campaigns" in Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop), 2024, Bangkok, Thailand

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

RÖTTGER, P., V. HOFMANN, V. PYATKIN, M. HINCK, H. KIRK, H. SCHUETZE, D. HOVY, "Political Compass or Spinning Arrow? Towards More Meaningful Evaluations for Values and Opinions in Large Language Models" in Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2024, Bangkok, Thailand

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

ATTANASIO, G., B. SAVOLDI, D. FUCCI, D. HOVY, "Twists, Humps, and Pebbles: Multilingual Speech Recognition Models Exhibit Gender Performance Gaps" in Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 2024, Miami, FL, United States of America

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

PLAZA DEL ARCO, F. M., D. NOZZA, D. HOVY, "Wisdom of Instruction-Tuned Language Model Crowds. Exploring Model Label Variation" in Proceedings of the 3rd Workshop on Perspectivist Approaches to NLP (NLPerspectives) @ LREC-COLING 2024, May 20-25, 2024, Torino, Italy

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

RÖTTGER, P., H. KIRK, B. VIDGEN, G. ATTANASIO, F. BIANCHI, D. HOVY, "XSTest: A Test Suite for Identifying Exaggerated Safety Behaviours in Large Language Models" in 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

### **"My Answer is C": First-Token Probabilities Do Not Match Text Answers in Instruction-Tuned Language Models**

WANG, X., B. MA, C. HU, L. WEBER-GENZEL, P. RÖTTGER, F. KREUTER, D. HOVY, B. PLANK, "'My Answer is C': First-Token Probabilities Do Not Match Text Answers in Instruction-Tuned Language Models" in Findings of the Association for Computational Linguistics ACL 2024, 2024, Bangkok, Thailand

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

HADA, R., A. EBRAHIMI FARD, S. SHUGARS, F. BIANCHI, P. ROSSINI, D. HOVY, R. TROMBLE, N. TINTAREV, "Beyond Digital "Echo Chambers": The Role of Viewpoint Diversity in Political Discussion" in WSDM '23: The Sixteenth ACM International Conference on Web Search and Data Mining, 27 February - 3 March, 2023, Singapore

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

HUNG, C.-C., A. LAUSCHER, D. HOVY, S. P. PONZETTO, G. GLAVAŠ, "Can demographic factors improve text classification? Revisiting demographic adaptation in the age of transformers" in Findings of the Association for Computational Linguistics: EACL 2023, May 2-6, 2023, Dubrovnik, Croatia

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

CERCAS CURRY, A., G. ATTANASIO, D. HOVY, D. NOZZA, "MilaNLP at SemEval-2023 Task 10: Ensembling Domain-Adapted and Regularized Pretrained Language Models for Robust Sexism Detection" in Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023), July 13-14, 2023, Toronto, Canada

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

ORLIKOWSKI, M., P. RÖTTGER, P. CIMIANO, D. HOVY, "The ecological fallacy in annotation: modeling human label variation goes beyond sociodemographics" in Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, July 9-14, 2023, Toronto, Canada

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

THEIL, K., D. HOVY, H. STUCKENSCHMIDT, "Top-down influence? Predicting CEO personality and risk impact from speech transcripts" in Proceedings of the Seventeenth International AAAI Conference on Web and Social Media, June 5–8, 2023, Limassol, Cyprus

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

LAUSCHER, A., D. NOZZA, A. CROWLEY, E. MILTERSEN, D. HOVY, "What about em? How commercial Machine Translation fails to handle (neo-)pronouns" in Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, July 9-14, 2023, Toronto, Canada

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

BIANCHI, F., S. HILLS, P. ROSSINI, D. HOVY, R. TROMBLE, N. TINTAREV, ""It's not just hate": a multi-dimensional perspective on detecting harmful speech online" in The 2022 Conference on Empirical Methods in Natural Language Processing, December 7-11, 2022, Abu Dhabi, United Arab Emirates

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

ATTANASIO, G., D. NOZZA, E. PASTOR, D. HOVY, "Benchmarking Post-Hoc Interpretability Approaches for Transformer-based Misogyny Detection" in Proceedings of NLP Power! The First Workshop on Efficient Benchmarking in NLP, May 22-27, 2022, Dublin, Ireland

### **Bridging fairness and environmental sustainability in natural language processing**

HESSENTHALER, M., E. STRUBELL, D. HOVY, A. LAUSCHER, "Bridging fairness and environmental sustainability in natural language processing" in The 2022 Conference on Empirical Methods in Natural Language Processing, December 7-11, 2022, Abu Dhabi, United Arab Emirates

### **Entropy-based Attention Regularization Frees Unintended Bias Mitigation from Lists**

ATTANASIO, G., D. NOZZA, D. HOVY, E. BARALIS, "Entropy-based Attention Regularization Frees Unintended Bias Mitigation from Lists" in Findings of the Association for Computational Linguistics: ACL 2022, May 22-27, 2022, Dublin, Ireland

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

FORNACIARI, T., A. UMA, M. POESIO, D. HOVY, "Hard and Soft Evaluation of NLP models with BOOtSTrap SAMpling - BooStSa" in Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, May 22-27, 2022, Dublin, Ireland

### **Language Invariant Properties in Natural Language Processing**

BIANCHI, F., D. NOZZA, D. HOVY, "Language Invariant Properties in Natural Language Processing" in Proceedings of NLP Power! The First Workshop on Efficient Benchmarking in NLP, May 22-27, 2022, Dublin, Ireland

### **Pipelines for Social Bias Testing of Large Language Models**

NOZZA, D., F. BIANCHI, D. HOVY, "Pipelines for Social Bias Testing of Large Language Models" in Proceedings of BigScience Episode #5 -- Workshop on Challenges & Perspectives in Creating Large Language Models, May 22-27, 2022, Dublin, Ireland

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

BIANCHI, F., V. CUTRONA, D. HOVY, "Twitter-demographer: a flow-based tool to enrich Twitter data" in The 2022 Conference on Empirical Methods in Natural Language Processing, December 7-11, 2022, Abu Dhabi, United Arab Emirates

### **XLM-EMO: Multilingual Emotion Prediction in Social Media Text**

BIANCHI, F., D. NOZZA, D. HOVY, "XLM-EMO: Multilingual Emotion Prediction in Social Media Text" in Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, May 22-27, 2022, Dublin, Ireland

---

## **Other**

### **IA & società**

HOVY, D., "IA & società", Eco, July 2024

### **AI e sport business: le regole del gioco sono cambiate**

HOVY, D., A. CIRRINCIONE, "AI e sport business: le regole del gioco sono cambiate", Capital, December 2023, pp. 40-43

### **La manifattura è diventata intelligente**

## Research Notes or short Articles in Journals

### **Viewpoint: Artificial Intelligence Accidents Waiting to Happen?**

BIANCHI, F., A. CERCAS CURRY, D. HOVY, "Viewpoint: Artificial Intelligence Accidents Waiting to Happen?", Journal of Artificial Intelligence Research (JAIR), 2023, vol. 76, pp. 193-199

---

## Articles upon Request/Invitation

### **Five sources of bias in natural language processing**

HOVY, D., S. PRABHUMOYE, "Five sources of bias in natural language processing", Language and Linguistics Compass, 2021, vol. 15, no. 8, pp. e12432