

Data Statements:  
Documenting the datasets used for  
training and testing natural language processing systems

Angelina McMillan-Major, Emily M. Bender, and Batya Friedman

Presented at: Scholarly Communication in Linguistics: Resource Workshop and Poster Session, LSA 2022, January 7, 2022

## References

- Bender, E. M., Friedman, B., and McMillan-Major, A. (2021a). A guide for writing data statements for natural language processing. [https://techpolicylab.uw.edu/wp-content/uploads/2021/11/Data\\_Statements\\_Guide\\_V2.pdf](https://techpolicylab.uw.edu/wp-content/uploads/2021/11/Data_Statements_Guide_V2.pdf).
- Bender, E. M. and Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics*, 6:587–604.
- Bender, E. M., Gebru, T., McMillan-Major, A., Shmitchell, S., and et al (2021b). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of FAccT 2021*.
- Blodgett, S. L., Barocas, S., Daumé III, H., and Wallach, H. (2020). Language (technology) is power: A critical survey of “bias” in NLP. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5476, Online. Association for Computational Linguistics.
- Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., and Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In Lee, D. D., Sugiyama, M., Luxburg, U. V., Guyon, I., and Garnett, R., editors, *Advances in Neural Information Processing Systems 29*, pages 4349–4357. Curran Associates, Inc.
- Caliskan, A., Bryson, J. J., and Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186.
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., III, H. D., and Crawford, K. (2021). Datasheets for datasets. *Commun. ACM*, 64(12):8692.
- Hovy, D. and Søgaard, A. (2015). Tagging performance correlates with author age. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 483–488. Association for Computational Linguistics.
- Jørgensen, A., Hovy, D., and Søgaard, A. (2015). Challenges of studying and processing dialects in social media. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 9–18. Association for Computational Linguistics.
- Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., and Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences*, 117(14):7684–7689.
- McMillan-Major, A., Bender, E. M., and Friedman, B. (forthcoming). Data statements: From concept to community practice.
- Speer, R. (2017). Conceptnet numberbatch 17.04: better, less-stereotyped word vectors. Blog post, <https://blog.conceptnet.io/2017/04/24/conceptnet-numberbatch-17-04-better-less-stereotyped-word-vectors/>, accessed 6 July 2017.
- Tatman, R. (2017). Gender and dialect bias in YouTube’s automatic captions. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, pages 53–59. Association for Computational Linguistics.
- Wassink, A. B. (2021). Uneven success: Racial bias in automatic speech recognition. Talk presented at the University of Michigan MLK Colloquium. Recording available at <https://youtu.be/CFKTxUmLByo>.
- Zhao, J., Wang, T., Yatskar, M., Ordonez, V., and Chang, K.-W. (2017). Men also like shopping: Reducing gender bias amplification using corpus-level constraints. In *Proceedings of the 2017 Conference on*

*Empirical Methods in Natural Language Processing*, pages 2941–2951. Association for Computational Linguistics.