

*exported from Zotero collection

** required papers in bold

- Ball-Burack, Ari, et al. “Differential Tweetment: Mitigating Racial Dialect Bias in Harmful Tweet Detection.” *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, ACM, 2021, pp. 116–28. DOI.org (Crossref), <https://doi.org/10.1145/3442188.3445875>.
- . “Differential Tweetment: Mitigating Racial Dialect Bias in Harmful Tweet Detection.” *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 2021, pp. 116–28. ACM Digital Library, <https://doi.org/10.1145/3442188.3445875>.
- . “Differential Tweetment: Mitigating Racial Dialect Bias in Harmful Tweet Detection.” *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, ACM, 2021, pp. 116–28. DOI.org (Crossref), <https://doi.org/10.1145/3442188.3445875>.
- Ball-Burack et al. - 2021 - Differential Tweetment Mitigating Racial Dialect .Pdf. <https://dl.acm.org/doi/pdf/10.1145/3442188.3445875>. Accessed 12 Nov. 2021.
- . <https://dl.acm.org/doi/pdf/10.1145/3442188.3445875>. Accessed 12 Nov. 2021.
- Belz, Anya, et al. “A Systematic Review of Reproducibility Research in Natural Language Processing.” *ArXiv:2103.07929 [Cs]*, Mar. 2021. *arXiv.org*, <http://arxiv.org/abs/2103.07929>.
- Bender, Emily M., et al. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜.” *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, ACM, 2021, pp. 610–23. DOI.org (Crossref), <https://doi.org/10.1145/3442188.3445922>.
- Bender, Emily M., and Batya Friedman. “Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science.” *Transactions of the Association for Computational Linguistics*, vol. 6, 2018, pp. 587–604. ACLWeb, https://doi.org/10.1162/tac1_a_00041.
- Benton, Adrian, et al. “Ethical Research Protocols for Social Media Health Research.” *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, Association for Computational Linguistics, 2017, pp. 94–102. ACLWeb, <https://doi.org/10.18653/v1/W17-1612>.
- Buchanan, Elizabeth A., and Charles M. Ess. “Internet Research Ethics and the Institutional Review Board: Current Practices and Issues.” *ACM SIGCAS Computers and Society*, vol. 39, no. 3, Dec. 2009, pp. 43–49. December 2009, <https://doi.org/10.1145/1713066.1713069>.
- Chang, Kai-Wei. *Tutorial: Bias and Fairness in Natural Language Processing*. <http://web.cs.ucla.edu/~kwchang/talks/emnlp19-fairnlp/>. Accessed 12 Nov. 2021.
- CS269 - Fairness, Accountability, and Transparency in Natural Language Processing. <https://uclanlp.github.io/CS269-Winter2020/>. Accessed 12 Nov. 2021.
- “Ethics FAQ.” *NAACL-HLT 2021*, <https://iuliatuirc-google.github.io/naacl-org/naacl-2021-website/ethics/faq/>. Accessed 12 Nov. 2021.
- Feigenbaum, Joan, et al. “Towards a Formal Model of Accountability.” *Proceedings of the 2011 New Security Paradigms Workshop*, Association for Computing Machinery, 2011, pp. 45–56. ACM Digital Library, <https://doi.org/10.1145/2073276.2073282>.
- “Final Call For Papers.” *NAACL-HLT 2021*, <https://iuliatuirc-google.github.io/naacl-org/naacl-2021-website/calls/papers/>. Accessed 12 Nov. 2021.
- Gebru, Timnit, et al. “Datasheets for Datasets.” *ArXiv:1803.09010 [Cs]*, Mar. 2020. *arXiv.org*, <http://arxiv.org/abs/1803.09010>.
- Henderson, Peter, et al. “Ethical Challenges in Data-Driven Dialogue Systems.” *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, ACM, 2018, pp. 123–29. DOI.org (Crossref), <https://doi.org/10.1145/3278721.3278777>.
- II, Alvin Grissom. 2016-Ml-Course. 2016. 2019. GitHub, <https://github.com/acgrissom/2016-ml-course/blob/1a8c446a78895d3b8a5e30eac59ed4b72d5c1aec/assignments/writing.md>.
- Jin, Zhijing, et al. “How Good Is NLP? A Sober Look at NLP Tasks through the Lens of Social Impact.” *ArXiv:2106.02359 [Cs]*, July 2021. *arXiv.org*, <http://arxiv.org/abs/2106.02359>.
- Linguistics 575: Ethics in NLP*. https://faculty.washington.edu/ebender/2019_575/scicomm.html. Accessed 12 Nov. 2021.
- Munger, Kevin. “Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment.” *Political Behavior*, vol. 39, no. 3, Sept. 2017, pp. 629–49. DOI.org (Crossref), <https://doi.org/10.1007/s11109-016-9373-5>.
- NAACL 2018 SRNLP Tutorial. https://docs.google.com/presentation/d/1HkKDbpph3RP9s3hdTEQ8rjy8PZDx4tSTW6hZily_yIc. Accessed 12 Nov. 2021.

*exported from Zotero collection

** required papers in bold

- . https://docs.google.com/presentation/d/1HkKDbpph3RP9s3hdTEQ8rjy8PZDx4tSTW6hZily_yIc. Accessed 12 Nov. 2021.
- Prabhumoye, Shrimai, et al. "Principled Frameworks for Evaluating Ethics in NLP Systems." *ArXiv:1906.06425 [Cs]*, June 2019. *arXiv.org*, <http://arxiv.org/abs/1906.06425>.
- "Reproducibility Checklist." *NAACL-HLT 2021*, <https://iuliaturc-google.github.io/naacl-org/naacl-2021-website/calls/reproducibility-checklist/>. Accessed 12 Nov. 2021.
- Santy, Sebastin, et al. "Use of Formal Ethical Reviews in NLP Literature: Historical Trends and Current Practices." *ArXiv:2106.01105 [Cs]*, June 2021. *arXiv.org*, <http://arxiv.org/abs/2106.01105>.**
- Shmueli, Boaz, et al. "Beyond Fair Pay: Ethical Implications of NLP Crowdsourcing." *ArXiv:2104.10097 [Cs]*, Apr. 2021. *arXiv.org*, <http://arxiv.org/abs/2104.10097>.
- Shuster, Evelyne. "Fifty Years Later: The Significance of the Nuremberg Code." *New England Journal of Medicine*, vol. 337, no. 20, Nov. 1997, pp. 1436–40. *DOI.org (Crossref)*, <https://doi.org/10.1056/NEJM199711133372006>.
- "Why Facebook Is Losing the War on Hate Speech in Myanmar." *Reuters*. *www.reuters.com*, <https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/>. Accessed 12 Nov. 2021.
- Williams, Matthew L., et al. "Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views, Online Context and Algorithmic Estimation." *Sociology*, vol. 51, no. 6, Dec. 2017, pp. 1149–68. *DOI.org (Crossref)*, <https://doi.org/10.1177/0038038517708140>.
- WWW 2019 SRNLP Tutorial Final*. <https://docs.google.com/presentation/d/1lePB796BFCwIjBtBtxkgopinL5vmI0GPs83EdJ0irqI>. Accessed 12 Nov. 2021.