

- Rodríguez-Sánchez, F., J. C. de Albornoz, L. Plaza, J. Gonzalo, P. Rosso, M. Comet, and T. Donoso. 2021. Overview of exist 2021: sexism identification in social networks. *Procesamiento del Lenguaje Natural*, 67(0):195–207.
- Sanguinetti, M., G. Comandini, E. di Nuovo, S. Frenda, M. Stranisci, C. Bosco, T. Caselli, V. Patti, and I. Russo. 2020. Haspeede 2 @ evalita2020: Overview of the evalita 2020 hate speech detection task. In V. Basile, D. Croce, M. Di Maro, and L. Passaro, editors, *Proceedings of the Seventh Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2020)*, volume 2765. CEUR Workshop Proceedings (CEUR-WS.org). Conference date: 17-12-2020.
- Sap, M., S. Gabriel, L. Qin, D. Jurafsky, N. A. Smith, and Y. Choi. 2020. Social bias frames: Reasoning about Social and power implications of language. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5477–5490, Online, July. Association for Computational Linguistics.
- Schmeisser-Nieto, W., M. Nofre, and M. Taulé. 2022. Criteria for the annotation of implicit stereotypes. In *Proceedings of the Language Resources and Evaluation Conference*, pages 753–762, Marseille, France, June. European Language Resources Association.
- Sánchez-Junquera, J., B. Chulvi, P. Rosso, and S. P. Ponzetto. 2021. How do you speak about immigrants? taxonomy and stereoisimmigrants dataset for identifying stereotypes about immigrants. *Applied Sciences*, 11(8).
- Tajfel, H. 1984. *Grupos humanos y categorías sociales*. Herder.
- Tajfel, H., A. A. Sheikh, and R. C. Gardner. 1964. Content of stereotypes and the inference of similarity between members of stereotyped groups. *Acta Psychologica*, 22(3):191–201.
- Taulé, M., A. Ariza, M. Nofre, E. Amigó, and P. Rosso. 2021. Overview of DETOXIS at IberLEF 2021: DETECTION of TOXicity in comments In Spanish. *Procesamiento del Lenguaje Natural*, 67(0):209–221.
- Uma, A., T. Fornaciari, A. Dumitrache, T. Miller, J. Chamberlain, B. Plank, E. Simpson, and M. Poesio. 2021. SemEval-2021 Task 12: Learning with Disagreements. In *Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021)*, pages 338–347, Online, August. Association for Computational Linguistics.
- Vázquez, J. M., V. P. Álvarez, C. T. Taybi, and P. P. Sánchez. 2022. I2C at IberLEF-2022 DETESTS task: Detection of Racist Stereotypes in Spanish Comments using UnderBagging and Transformers. In *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2022)*. CEUR Workshop Proceedings, CEUR-WS.org.

A Appendix: Co-occurrence of Stereotype Categories within a sentence

This appendix provides a heatmap of the co-occurrence of stereotype categories within a sentence to visually spot those categories that are used together more often (see Figure 5).

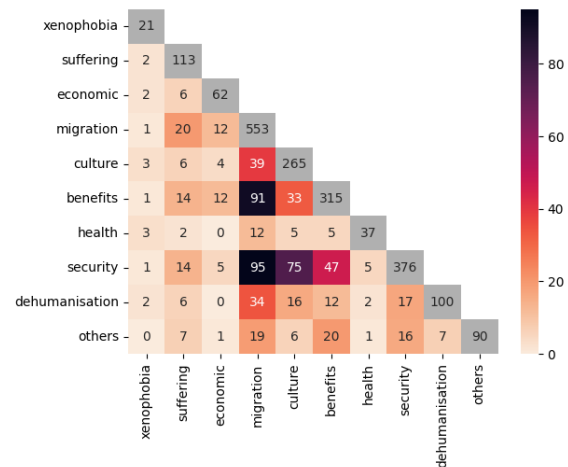


Figure 5: Heatmap representation of the sentence-level co-occurrence of stereotype categories with the occurrence count of each category coloured in gray.