

# FAIR AND PRIVATE DATA PREPROCESSING THROUGH MICROAGGREGATION

CARLOS GONZÁLEZ

*Newcastle University*

e-mail: carlos.gonzalez@ncl.ac.uk

Privacy protection for personal data and fairness in automated decisions are fundamental requirements for Responsible Machine Learning [3]. Both may be enforced through data preprocessing and share a common target: data should remain useful for a task, while becoming uninformative of the sensitive information. The intrinsic connection between privacy and fairness implies that modifications performed to guarantee one of these goals, may have an effect on the other, e.g., hiding a sensitive attribute from a classification algorithm might prevent a biased decision rule having such attribute as a criterion. In this talk, we present FAIR-MDAV [1], a fairness-and-privacy correcting mechanism based on the MDAV clustering algorithm [2]. This work resides at the intersection of algorithmic fairness and privacy: we show how the two goals are compatible and may be simultaneously achieved, with a small loss in predictive performance. Our results are competitive with both state-of-the-art fairness correcting algorithms and hybrid privacy-fairness methods. Experiments were performed on three widely used benchmark datasets: *Adult Income*, *COMPAS* and *German Credit*.

## References

- [1] C. González, J. Salas, D. Megías, and P. Missier. Fair and Private Data Preprocessing through Microaggregation. *ACM Transactions on Knowledge Discovery from Data* 18.3 (2023): 1–24.
- [2] Domingo-Ferrer, Josep, and Vicenç Torra. "Ordinal, continuous and heterogeneous k-anonymity through microaggregation." *Data Mining and Knowledge Discovery* 11 (2005): 195-212.
- [3] A. Chouldechova and A. Roth. A snapshot of the frontiers of fairness in machine learning. *Communications of the ACM*, 63(5):82–89, 2020.