Contribution ID: **70** Type: **not specified**

Exploring Hidden Societal Biases in Twitter Cascades: A Sociophysics Study

Human biases influence behavior and society, sometimes leading to discrimination and poor judgment. While algorithms were initially thought to be free from human biases, it's now understood that they can amplify existing biases, especially when trained on human-generated data. To address this, methods for identifying and mitigating biases in machine learning (ML) algorithms have been developed, focusing on auditing training datasets or model outputs. However, these tools struggle to identify unknown biases. Our project aims to uncover hidden biases using statistical anomaly detection methods, focusing on social media, particularly Twitter. We'll analise Twitter data, including user information like gender, nationality, and follower count, to detect patterns of biased information sharing. By examining the growth model parameters of Twitter cascades (sequences of retweets), we aim to find statistical differences that indicate bias. This project offers practical experience in handling large datasets and statistical models, and potentially, new insights into identifying online biases.

Author: SILVA, Tomás

Presenter: SILVA, Tomás