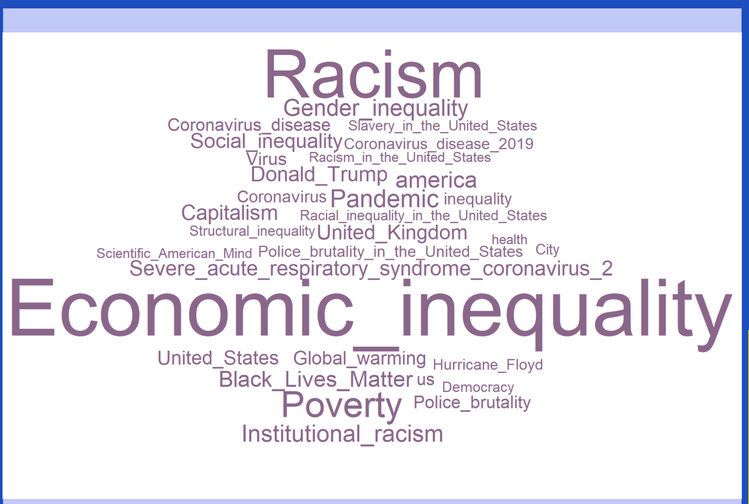
Visualizing Social Debates on Inequality from a Large Social Media Knowledge Graph

Alkım Belevi (Sociology), Evy Beijen (Mathematics/Religious Studies),
Mona Borth (Mathematics), Celeste Tacconi (Linguistics)

Research Question

What are some effective approaches to visualizing patterns in the data contained in the OKG on tweets discussing inequality (Blin et al., 2023) in order to better understand the discourse about inequality on social media?



Word cloud representing the most frequently mentioned entities in the dataset (>50 times)

Introduction

- Existing strategies for analyzing social media discourse often fail to capture fine-grained opinions and intricacies.
- Blin et al. (2023) structure a sample of Twitter data on inequality into an observatory knowledge graph (OKG).
- We visualize data from the OKG, bridging the gap between complex data and interpretable information, revealing patterns in narratives on inequality.

Challenges

- Computer science challenge: Finding suitable methods to visualize data from a complex large-scale knowledge graph.
- Social science challenge: Enabling social science researchers to access the insights from the OKG that are otherwise hard to obtain due to technical complexity.

Limitations

- Small data set: limited generalizability of results
- Sampling bias: only tweets containing the words “inequality” or “inequalities”
- Dataset not containing the full text of tweets
- Partly inaccurate entity linking
- Geographical locations not normalized
- Small participant pool for user study.

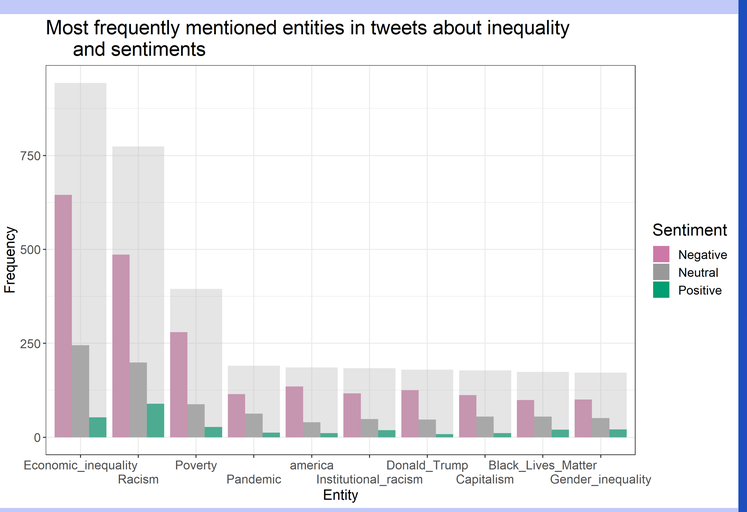
Dataset & Method

- The data is based on 62 015 tweets from May 30th to August 27th, 2020, containing the keywords “inequality” or “inequalities”.
- The OKG contains metadata and linguistic information about the tweets.
- We explored the data and loaded it into R through SPARQL queries.
- We created visualizations, mainly using R.
- Evaluation method: User study with two political science students and two of the creators of the OKG .

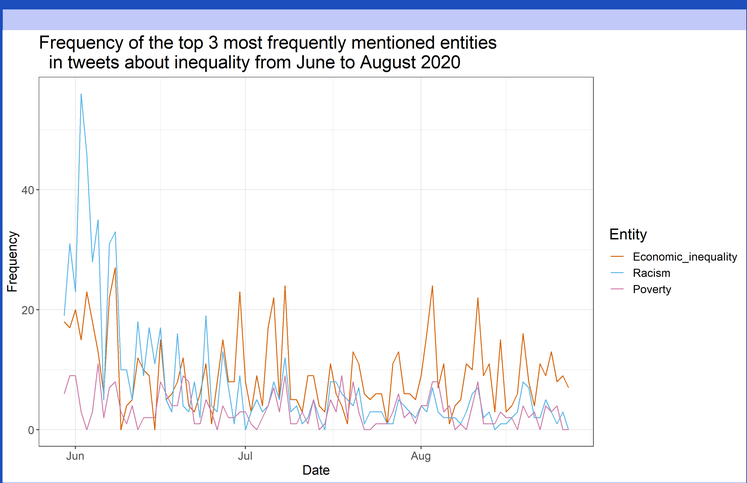
Results

- Product: User-friendly, accessible and inclusive gallery of various types of visualizations with accompanying text and code.
- User study results:
  - useful for **understanding social media discourse on inequality**
  - missing contextual information
  - explain the data selection process
  - incorporate clearer, more readable labels
  - preference for interactivity.
- The user study results were partly incorporated.

Entity: automatically recognized object, event, situation or concept

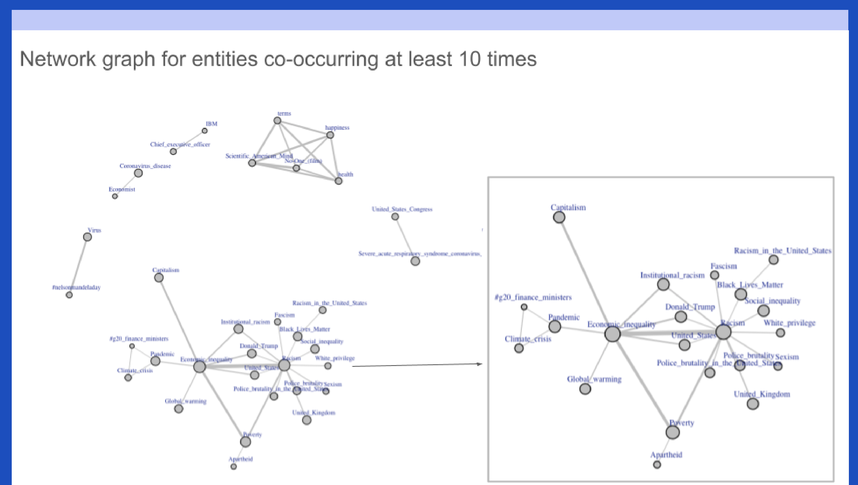


Most frequently mentioned entities, including sentiments

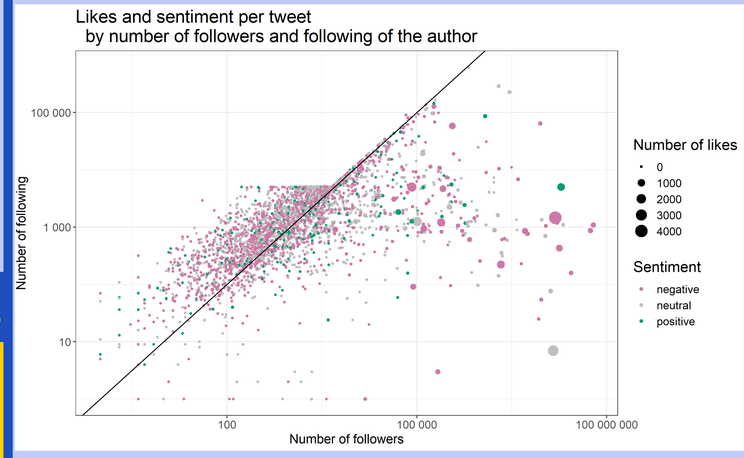


Frequency of top 3 most frequently mentioned entities

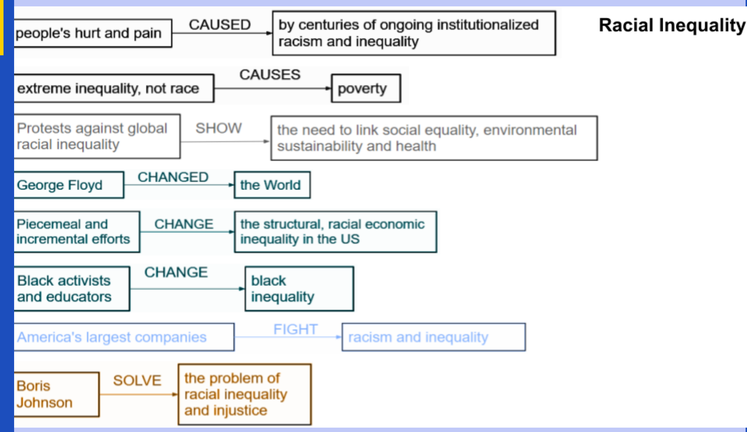
Knowledge Graph → **SPARQL Queries** → **Visualizations in R**



Network graph for co-occurring entities (>10 times)



Likes and sentiment per tweet by author’s number of followers and following



Claims made by Twitter users about racial inequality

SCAN ME