

**SICSS Edinburgh**



THE UNIVERSITY  
*of* EDINBURGH

# **CSS in Practice**

# **Global Analysis**

**Walid Magdy**

29-May-2025

# Task/Purpose

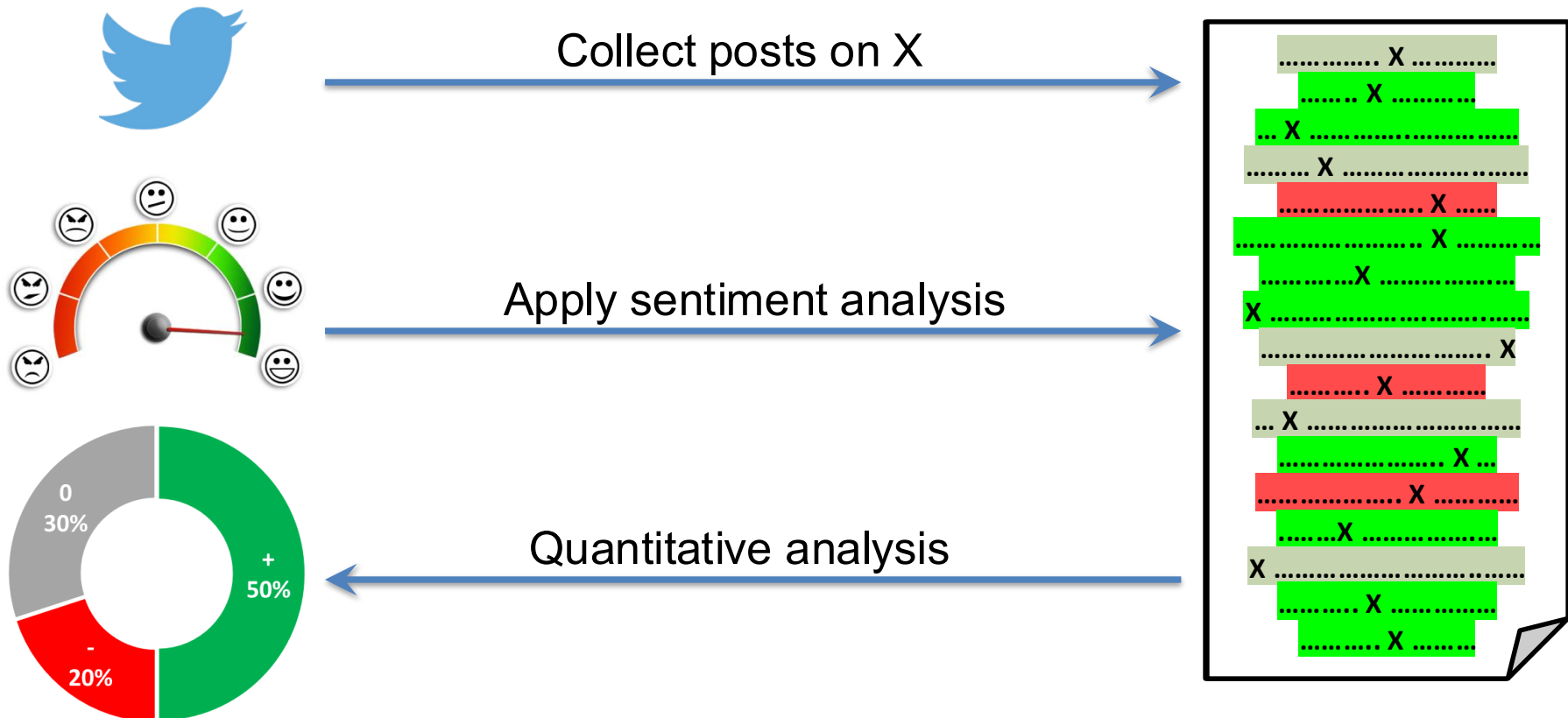
- Event, Phenomena, Topic  
→ understand what is going on (social science RQ)
- Important aspects:  
Trends, change over time
- Interest → **Global Analysis**
- Method → Collect data, label (classify), analyse
- Examples:
  - Elections, polarization, responds after major event ... etc.

# Outlines

- Examples of using CSS for global analysis
- 3 Example Studies
- No technical details (ask if you need details)
- Sharing main methodology
- Topics: might be sensitive!!

# Important Note: Using Sentiment to Measure Support

# How People Think about Topic X?

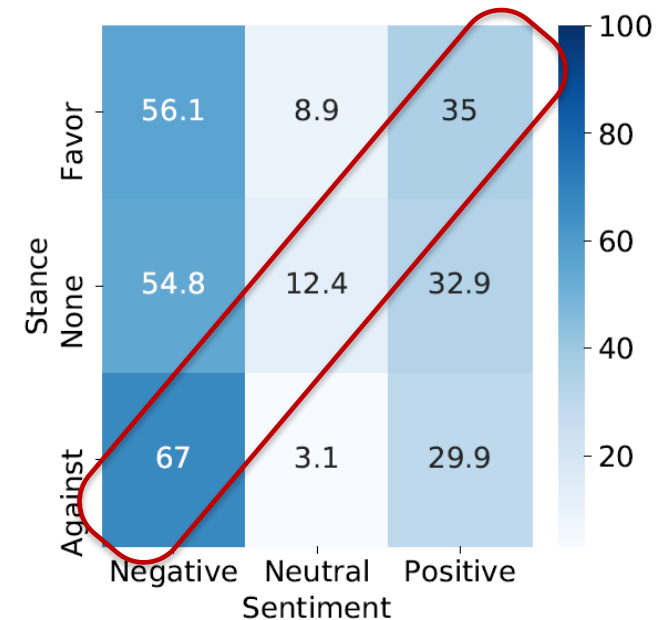


# Common Practice

- How people think about **X** (e.g. ~~iPhone 16~~, ~~SIGSS~~)?
- Collect posts mentioning **X**
- Use sentiment analyser to check posts polarity
- Result: 50% positive, 20% negative, 30% neutral
- Conclusion: 50% of posts ~~likes~~ **X**, while 20% are ~~negative to~~ it.  
  supports  
  against
- Can we use **sentiment** to measure **support** (stance)?

# Sentiment vs Stance

- Sentiment = emotion polarity in text
- Stance = position towards a given topic
- SemEval topics labelled independently on: Sentiment & Stance
- Measure agreement
- **No**, you cannot use sentiment to measure support!
- e.g.: “I am sad that Harris lost 😞”  
“*Trump is visiting Florida #MAGA*”



# Lesson

- **Sentiment  $\neq$  Stance**
- **When measuring support on a given topic/entity, use a stance classifier not a sentiment analyser**

## Ref:

- Aldayel A. and W. Magdy. Assessing Sentiment of the Expressed Stance on Social Media. SocInfo 2019



# Large data + Automatic Labeling

# People Changing Opinion?

# The Egyptian Military Intervention

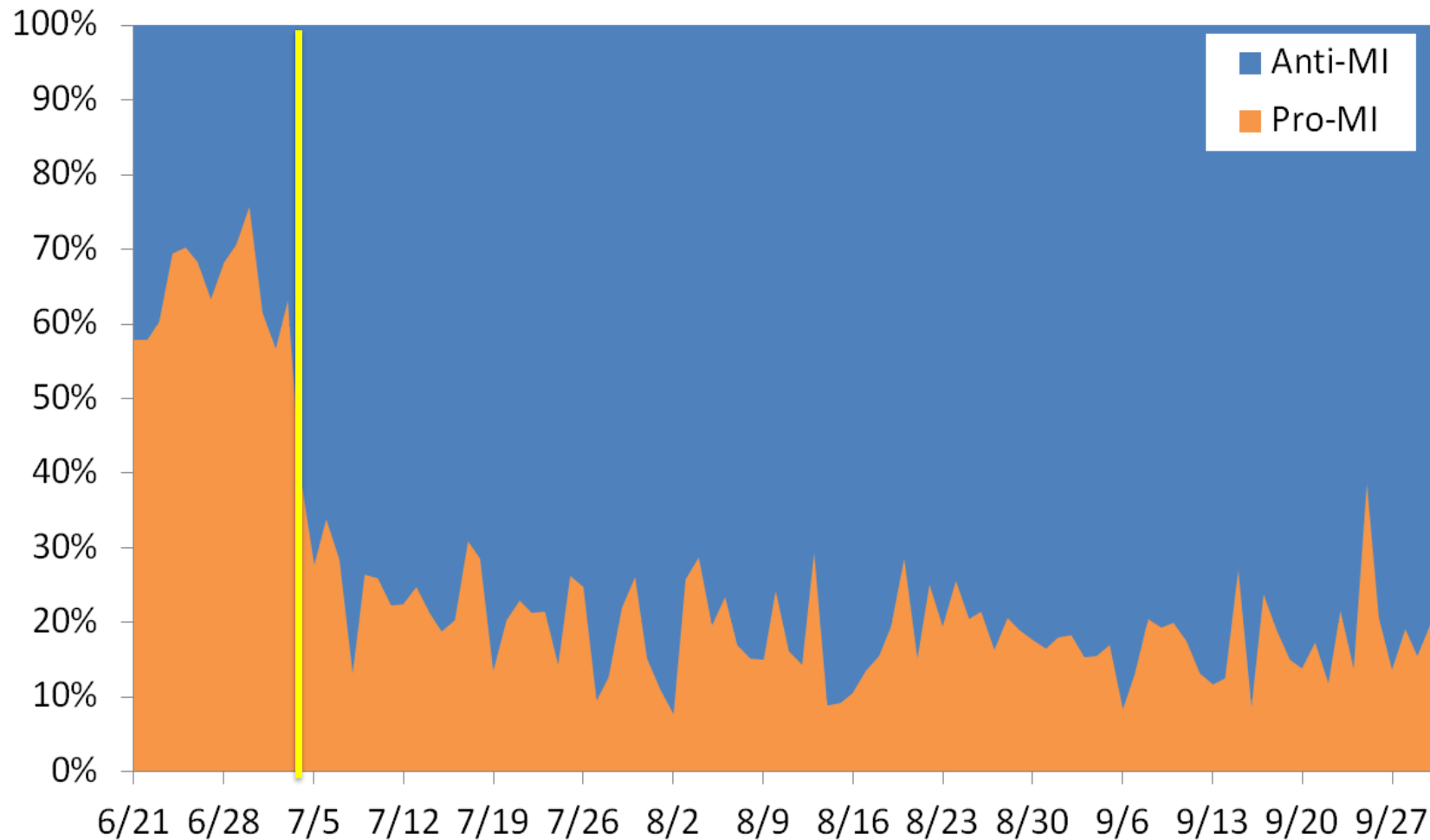


- 30 June 2013: large demonstration in Egypt against Morsi
- 3 July 2013: Military ousted Morsi
- 5 July-13 Aug: Large Sit-in against military coup
- 14 Aug: Army ends Sit-in by force, while hundreds killed

# Study

- Shift in trending tweets about Egypt was noticed since 3/7/2013
- **Data Collection:**
  - 6M tweets on Egypt → 21 July 2013 – 30 Sep 2013
- **Label tweets:**
  - Pro/Anti military intervention
  - Sample 1000 → train classifier → label the rest.
- **Classifier:**
  - Trained SVM binary classifier → Pro/Anti military intervention
  - Accuracy: 85% (on the tweet level)
  - Label all tweets on topic using the classifier

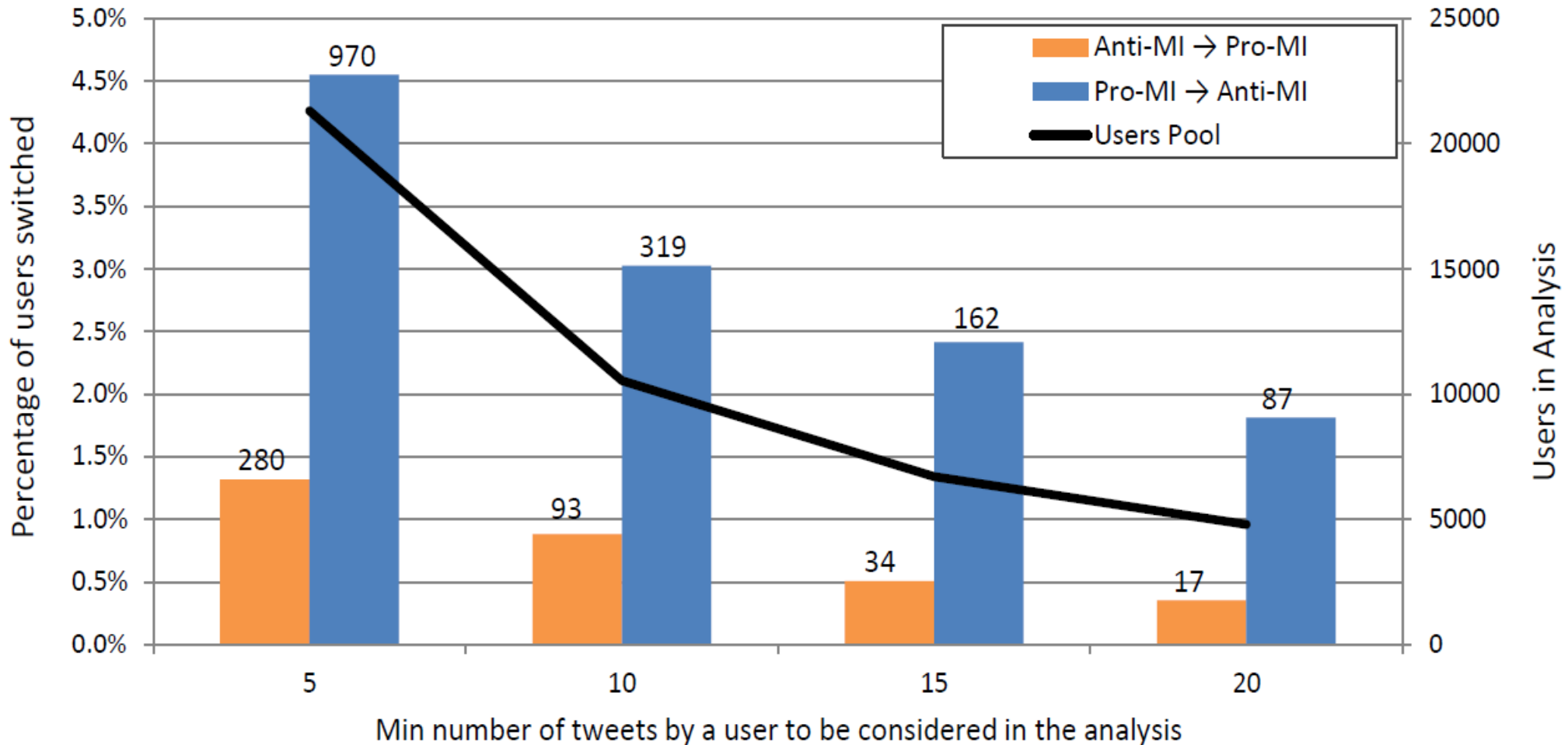
# Volume of Tweets Pro/Anti-MI



# On the User Level

- Identified 22K Twitter users with >5 tweets on topic
- Examine change in tweets stance over time
- Three confidence levels:
  - Users with 5+ tweets: 22K users
  - Users with 10+ tweets: 11K users
  - Users with 20+ tweets: 5K users
- Changes:
  - Pro-MI → Anti-MI
  - Anti-MI → Pro-MI

# % of Users Changing Opinion



# Methodology

- Data collection
- Sample labeling → classifier → label the rest
- Timeline analysis on:
  - Post level (trends) → can be misleading
  - User level → more indicative
- Lesson: It is not easy to have someone switching political belief

## Ref:

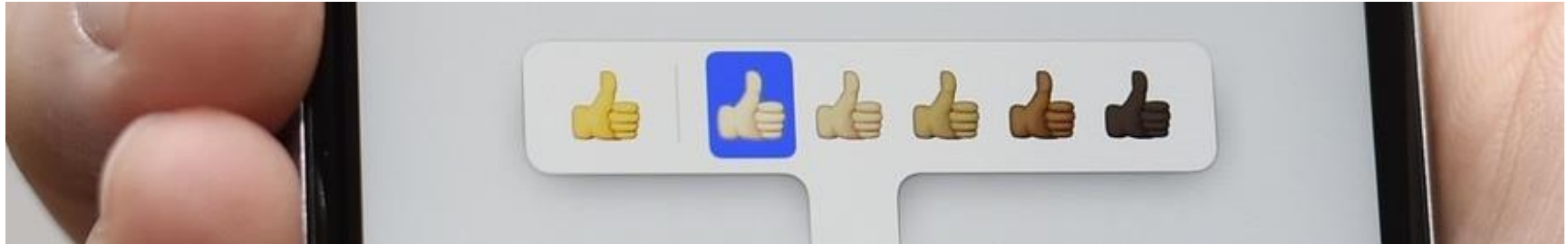
Borge-Holthoefer J., W. Magdy, K. Darwish, and I. Weber. Content and Network Dynamics Behind Egyptian Political Polarization on Twitter. *CSCW 2015*



# Large Data + Manual Labeling

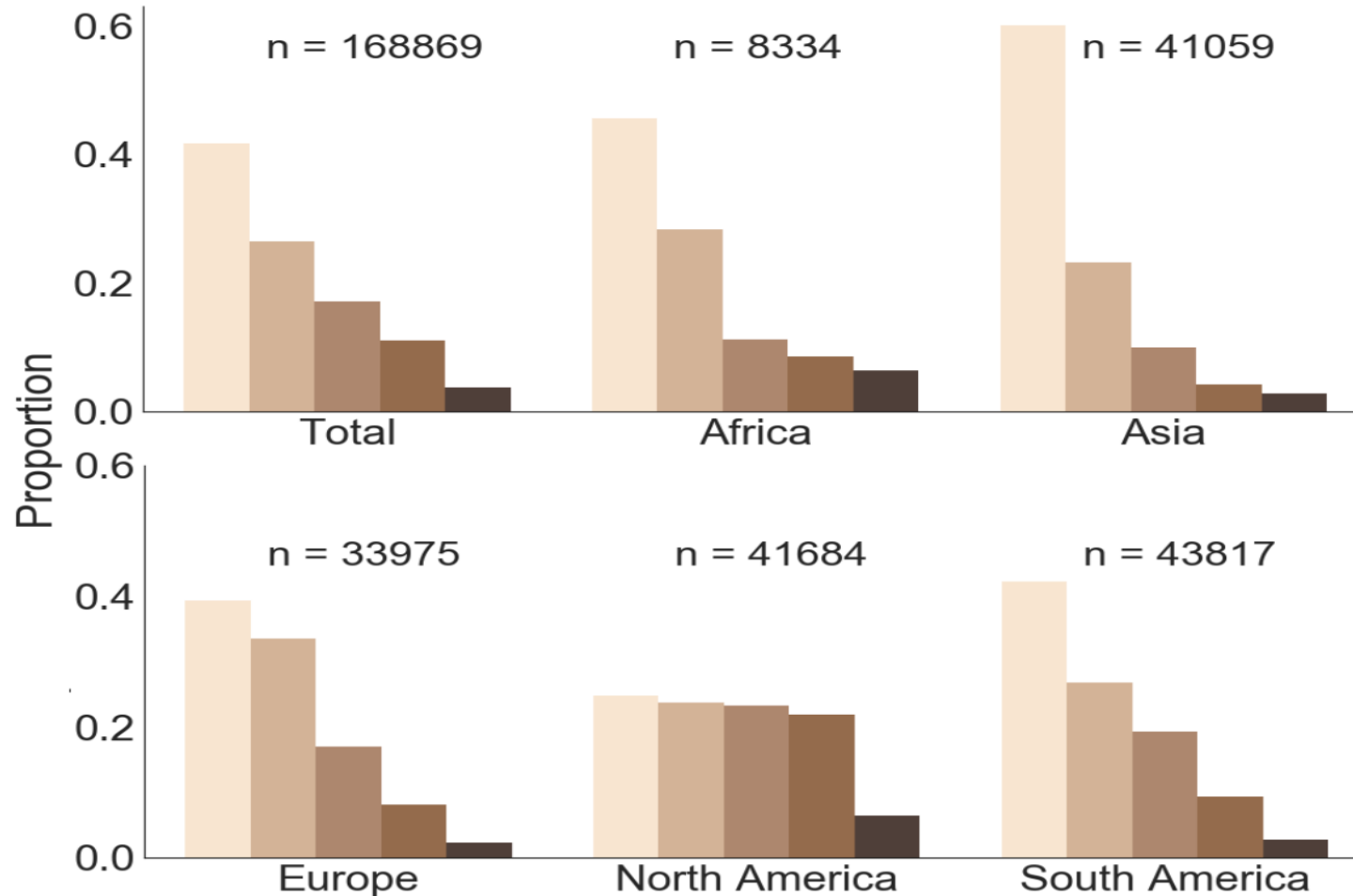
# General Usage of Features on Social Media

# Usage of Emoji skin-modifier

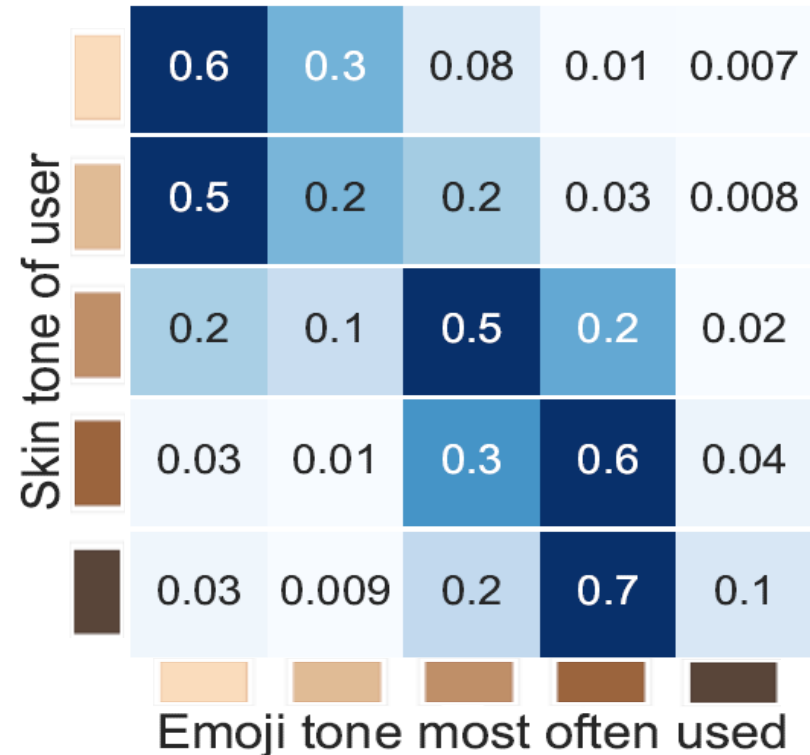
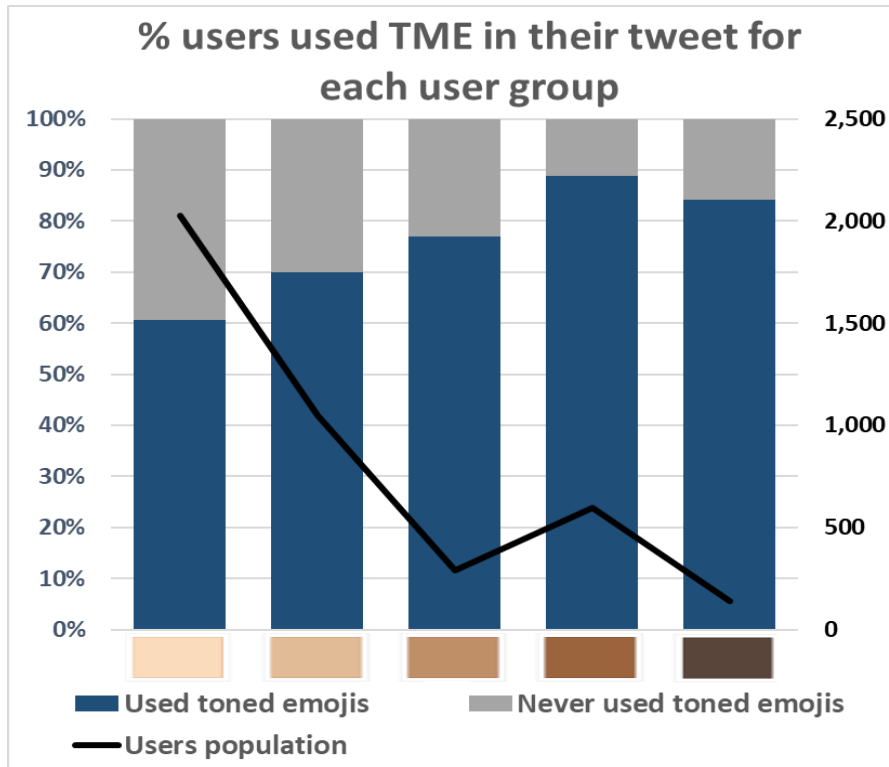


- Collection: **>1 billion** tweets from around the world
- Detect tweets with skin-modifiable emoji
- Check how skin color is modified by location
- Select **4,100** users (with their 7M tweets) who used these emoji
- Compare user skin color (manually annotated) to used color in emoji

# Emoji Skin-Tone by Region



# Who modifies the color? How?

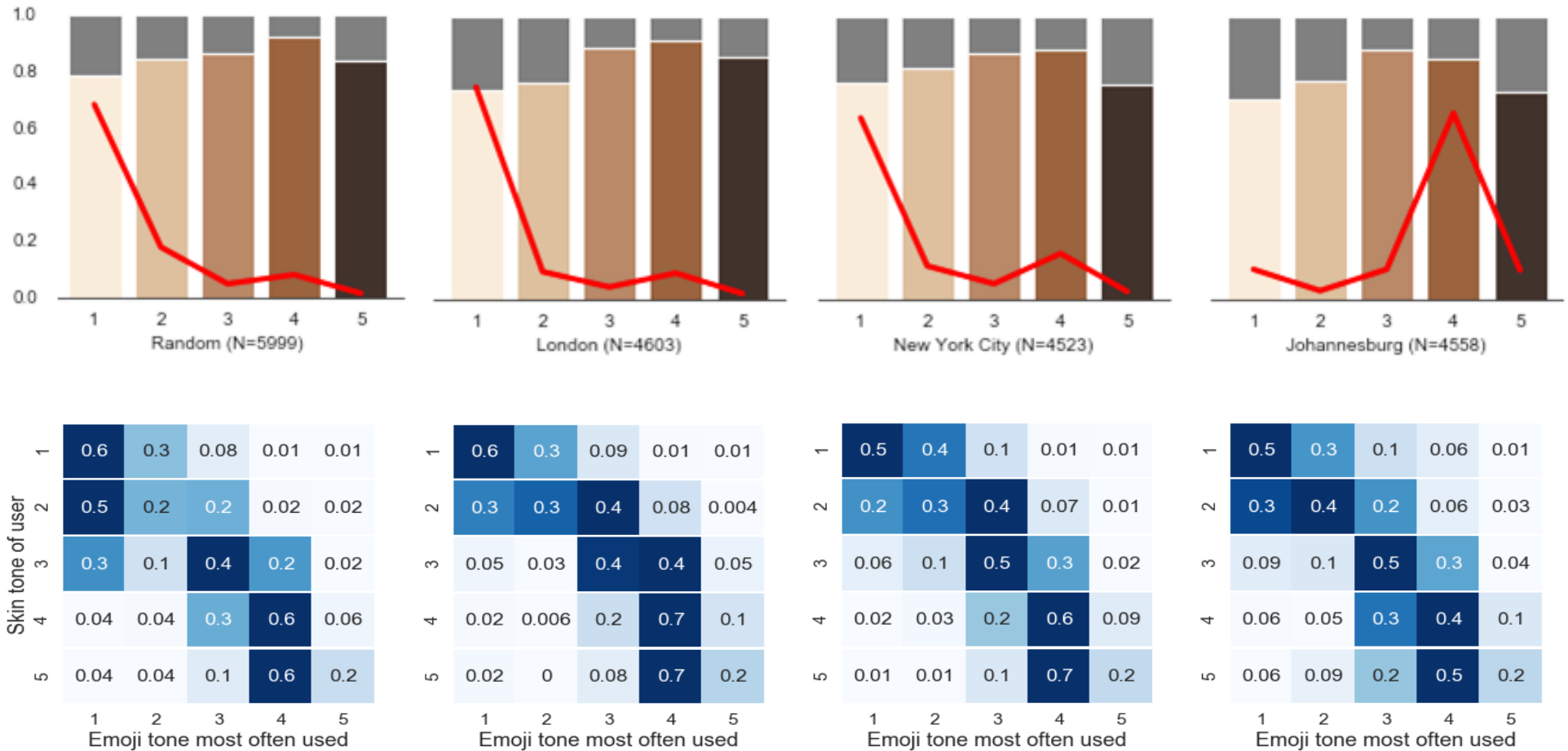


- Darker-skin users are more keen to modify skin-tone of emoji
- The majority of users use skin tones similar to their skin

# Generalizability of Findings

- Our sample is random set of users from all over the world!
- Are these findings consistent over multiple samples/locations?
- Labeled a new random set of ~6000 accounts
- Labeled additional accounts, but from specific locations:
  - **London** (~4600 accounts)
  - **New York City** (~4500 accounts)
  - **Johannesburg** (~4500 accounts)
- Repeated the analysis

# Generalizability of Findings



# Methodology

- Large data collection
- Manual labeling for users (ethical consideration!)
- Comparative analysis of usage

## Ref:

- Robertson A., W. Magdy, S. Goldwater. Self-Representation on Twitter Using Emoji Skin Color Modifiers. *ICWSM 2018*
- Robertson A., W. Magdy, S. Goldwater. Emoji Skin Tone Modifiers: Analyzing Variation in Usage on Social Media. *ACM Transaction of Social Computing*



# No/Limited Data!

# Facebook Content Moderation

## The Palestine/Israel Conflict

# About

**Gaza strip**



**The  
West Bank**

**Sheikh  
Jarrah**

# May 2021, The Event

## Sheikh Jarrah: Israeli police storm Palestinian protest over Jerusalem evictions

Israeli special forces use batons and mounted horses to disperse Palestinian and international activists protesting in solidarity with families facing eviction



Israeli forces detain a man after storming Palestinian houses whose families face eviction in the Sheikh Jarrah neighbourhood of East Jerusalem, 4 May 2021 (AFP)



facebook

We Removed Something You Posted

We removed the post below because it doesn't follow the Facebook Community Standards:

# Research Questions

1. Do deleted Arabic posts violate FB's community standards (FBCS)?
2. How do Arabs see the moderation process on these posts?
3. What is the content of these posts?

# Samples of the Deleted Posts

- Submission form + Survey



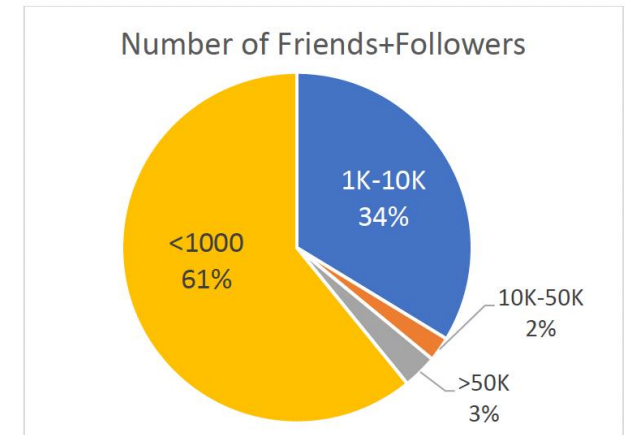
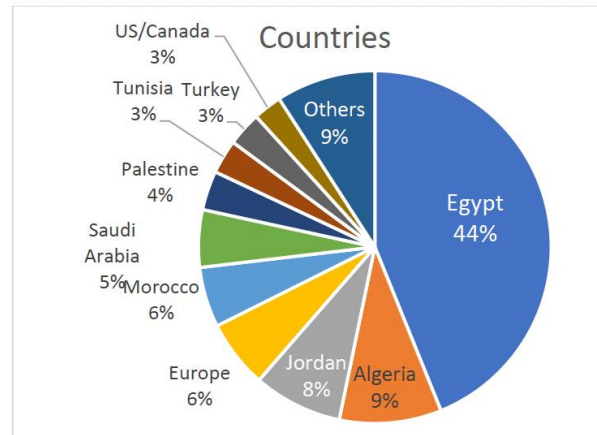
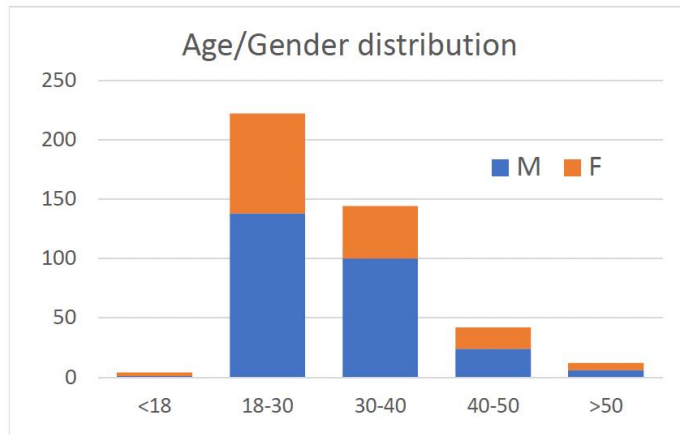
List of deleted Arabic posts

- Submit the deleted post (copy/paste)
- Action taken by FB + your opinion
- Manually labeled by topic

**448**  
**Posts**



# Collected Deleted Posts - Authors



Do you think FB is biased?

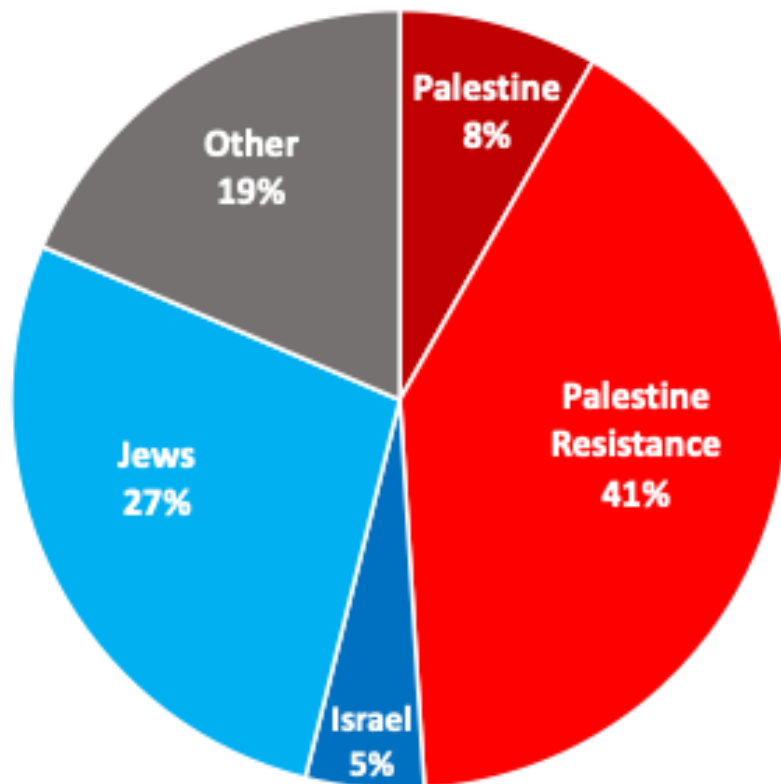
**93%**

FB is highly biased

**448  
Posts**

# Collected Deleted Posts - Topics




Topics of the collected posts



Theme	Description
Palestine	Posts showing support to Palestine and Palestinians
Israel	Posts about Israel, its cities, and Israelis
Palestine resistance	Posts supporting Palestine resistance, Hamas, Qassam, or their figures
Jews	Posts about Jews or Zionists (usually against them)
Others	Posts about other countries, suicide, men, women, and others religious groups



# Annotation

- Hire Arab annotators on  **appen**
- *Claim*: A social media platform!
- Task:
  - Check if a post violates guidelines 
  - Specify which of the 24 aspects violated
  - According to personal opinion, would you remove?
- **Quality Control!** 
- **10** annotators per post

**106**  
**Annotators**  
(6 Arab countries)

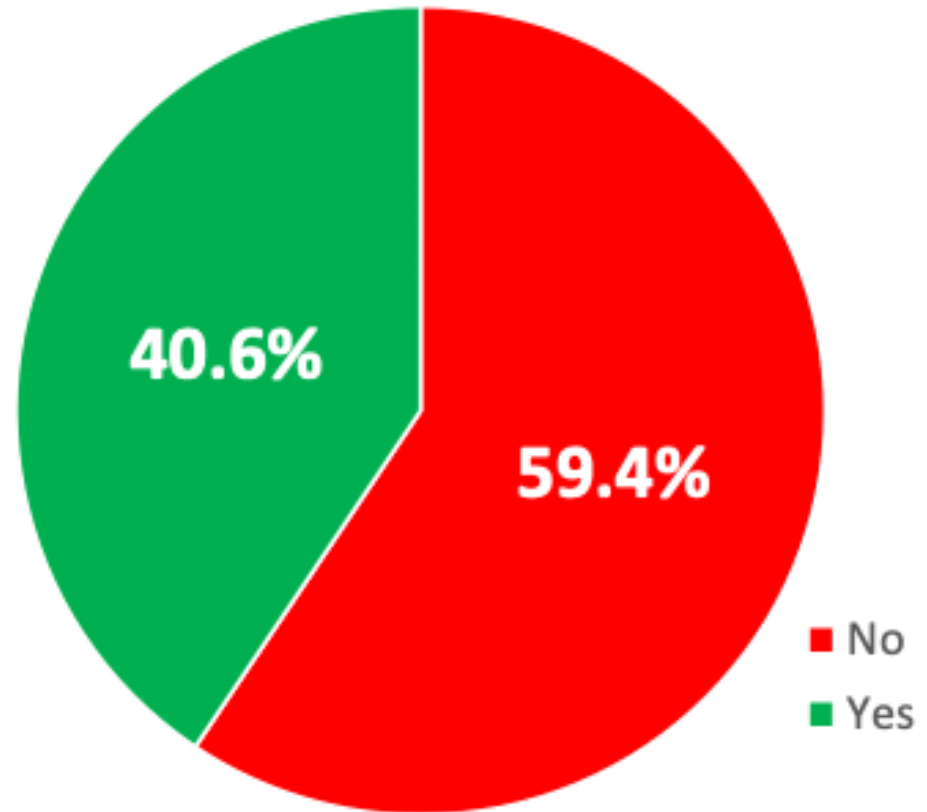
# Annotations – Does it violate FBCS?

Does the post violate the guidelines?

448 posts



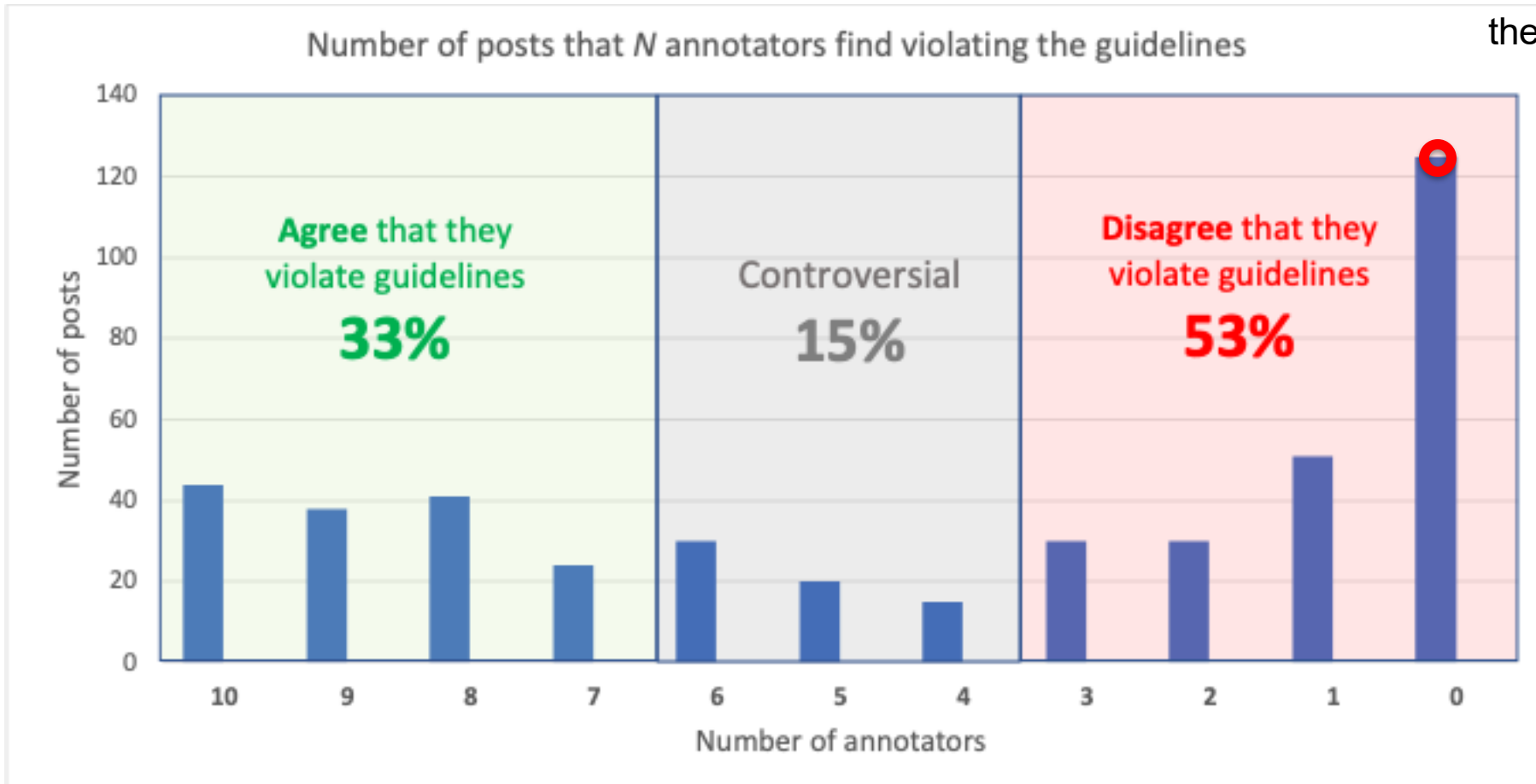
4480 annotations



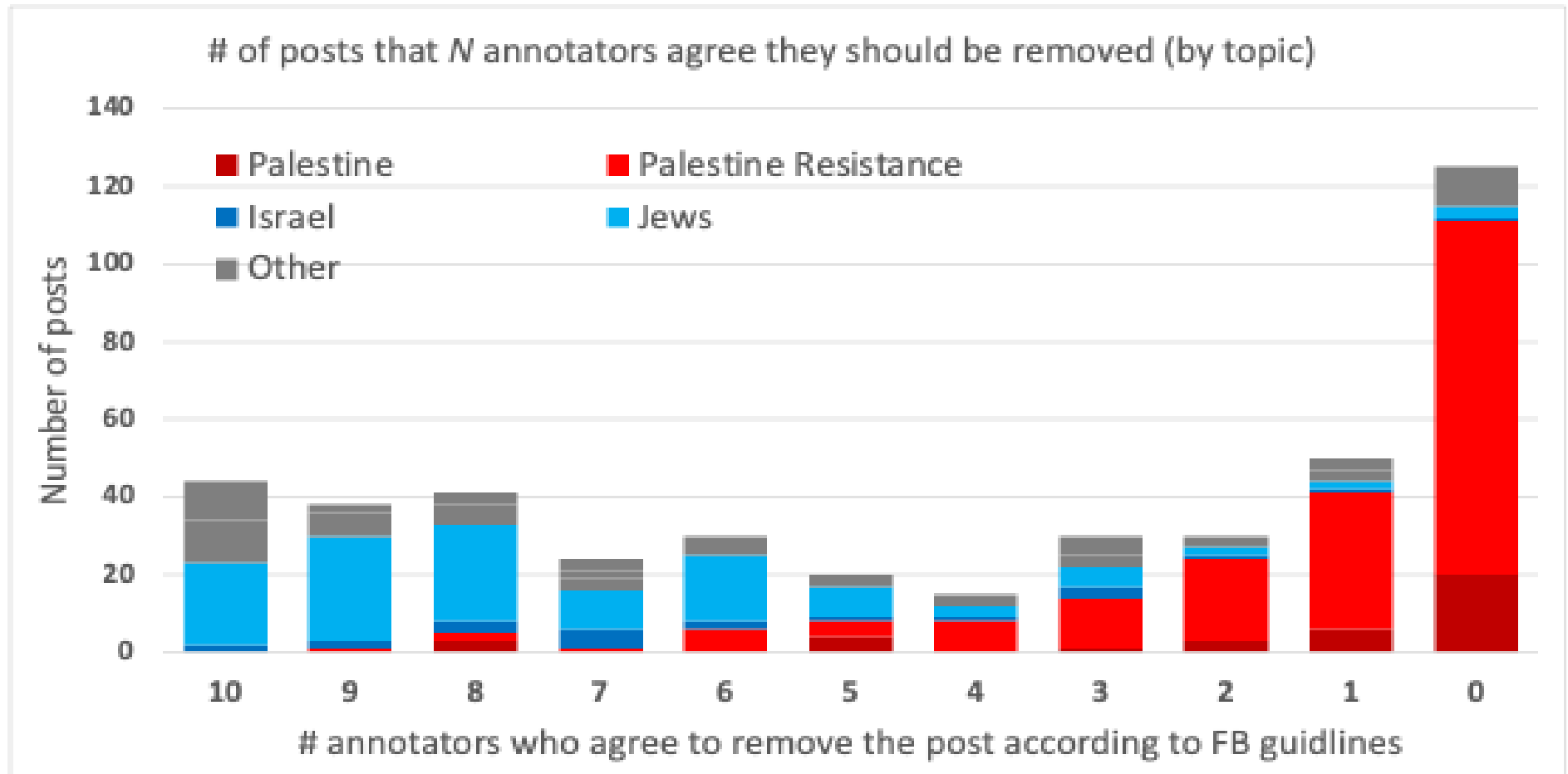
# Annotations – N/10 agree it violates?

**28%**

not a single  
annotator find  
them violate  
FBCS



# Annotations – by Topic?



# Is Facebook biased?

- Hate-speech, self-harm, ... etc.
- Political stance??!
- How to interpret guidelines?
  - Resistance vs terrorism!
- Underrepresented communities!
- Conflicts!
- Is it a double-standard?

# Methodology

- Sometimes data is challenging to collect
- Survey can be a solution
- Manual labeling → Annotation is the core of the study
- Smaller set of data + clear annotation methodology  
→ Solid study

## Ref:

Magdy W., H. Hussain, J. Salminen. Who should set the Standards? Analysing Censored Arabic Content on Facebook during the Palestine-Israel Conflict. **CHI 2025**

# Summary

- Many ways to analyse an event/topic globally using CSS
- Data:
  - Available on large scale (e.g. social media, logs)
  - Scarce: different methods for collection (surveys) or analysis (qualitative)
- Labeling strategy:
  - Automatic
  - Manual
  - Annotation can be the core of the study
- Labeling strategy:
  - Post level
  - User level
- Main tips:
  - Plan ahead + understand your limitations
  - Do NOT over generalize your findings

# Next

- Analyse users' behavior (on the individual level)
- Prediction something about the user



# Thank You

[wmagdy@inf.ed.ac.uk](mailto:wmagdy@inf.ed.ac.uk)

 [@walid\\_magdy](https://twitter.com/walid_magdy)

<http://homepages.inf.ed.ac.uk/wmagdy/>



 [@SMASH\\_Edin](https://twitter.com/SMASH_Edin)

<https://smash.inf.ed.ac.uk/>