

# Comparing aspects of usability in Mastodon and Twitter



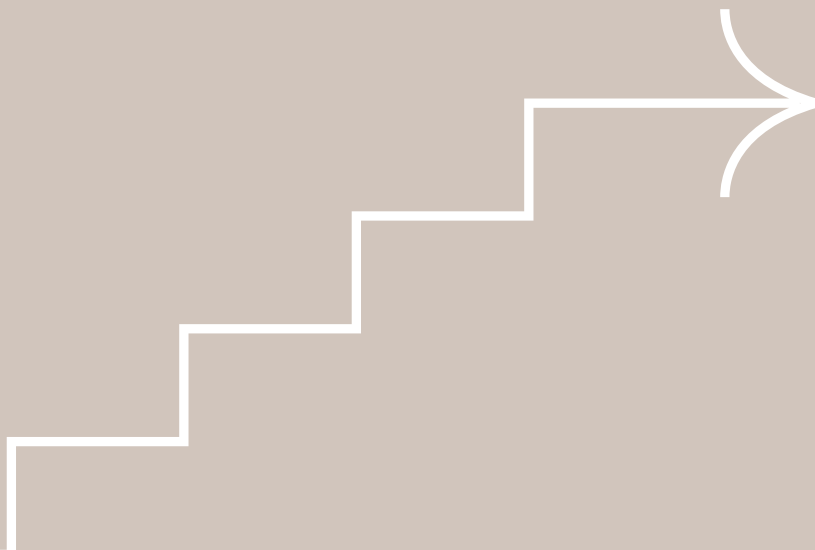
Data and Algorithmic governance

## **PRESENTED BY:**

Matthias Zöttl,  
Leonie Knill

## **PRESENTED TO:**

Sabrina Kirrane

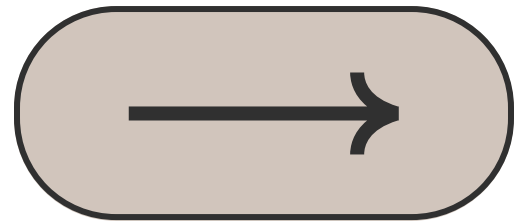


# ARE THERE DIFFERENCES IN USABILITY ON CENTRALIZED VS DECENTRALIZED PLATFORMS?

Quantitative analysis with focus on **user sentiment**

---

Qualitative analysis with focus on **polarizing content, bias / echo chambers and harmful content**



# Quantitative Analysis

Exploring usability in Mastodon vs. Twitter using  
sentiment analysis (Vader, BERT)

## Mastodon <> Twitter

### **USER SENTIMENT IN BITCOIN DISCUSSIONS:**

A TOPIC THAT IS SENSITIVE WITH  
REGARDS TO THE CENTRALIZED VS  
DECENTRALIZED DEBATE

### **DATA**

**Timeframe:** 05/02/21 - 23/06/21

**Hashtags:** #btc #bitcoin

**Mastodon:** 9130k toots

**Twitter:** 200k tweets





**vs**



# VADER

Valence Aware Dictionary for Sentiment Reasoning



Rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media

# roBERTa

Robustly Optimized Bidirectional Encoder Representations from Transformers Approach



Developed by Facebook AI, RoBERTa builds upon Google's BERT and can be used for a variety of NLP tasks

```

preprocess_tweet(tweet):
# Lowercasing
tweet = tweet.lower()

# Remove @usernames
tweet = re.sub('@[\w]+', '', tweet)

# Remove URLs
tweet = re.sub(r"http\S+|www\S+|https\S+", '', tweet, flags=re.MULTI

# Removing special characters and symbols except emojis
tweet = re.sub(
r"^[^a-zA-Z\s\U0001F300-\U0001F5FF\U0001F600-\U0001F64F\U0001F680-\U
r"\U0001F900-\U0001F9FF\U0001FA70-\U0001FAFF\U00002702-\U000027B0]+

# Lemmatization
lemmatizer = WordNetLemmatizer()
tokens = word_tokenize(tweet)
lemmatized_tokens = [lemmatizer.lemmatize(token) for token in token

# Removing stopwords
stop_words = set(stopwords.words("english"))
filtered_tokens = [token for token in lemmatized_tokens if token no

# Joining tokens
preprocessed_tweet = " ".join(filtered_tokens)

return preprocessed_tweet

```



transform to lowercase



remove @usernames



remove Links/Urls



removing non-alphanumeric  
characters except smileys 😊



Lemmatization - reducing words to their  
base form



remove stop words



Tokenization



transform to lowercase



remove @usernames



remove Links/Urls



removing non-alphanumeric  
characters except smileys 😊



Lemmatization - reducing words to their  
base form



remove stop words



Tokenization





# VADER

Valence Aware Dictionary for Sentiment Reasoning



Hi. #BTC pyramid need new suckers to pay the energy bill : Around \$20 millions daily. Grid operators cannot be paid in crypto. They need your fiat. Please send your life savings to one of the private exchange platforms and receive chocolate coins in exchange. Thanks in advance

9:30 PM • Frb 5, 2022



# roBERTa

Robustly Optimized Bidirectional Encoder Representations from Transformers Approach





**VADER**

Valence Aware Dictionary for Sentiment Reasoning



**Account**   
@usesrname



Hi. #BTC pyramid need new suckers to pay the energy bill :  
Around \$20 millions daily. Grid operators cannot be paid in crypto.  
They need your fiat. Please send your life savings to one of the  
private exchange platforms and receive chocolate coins in  
exchange. Thanks in advance

9:30 PM • Frb 5, 2022



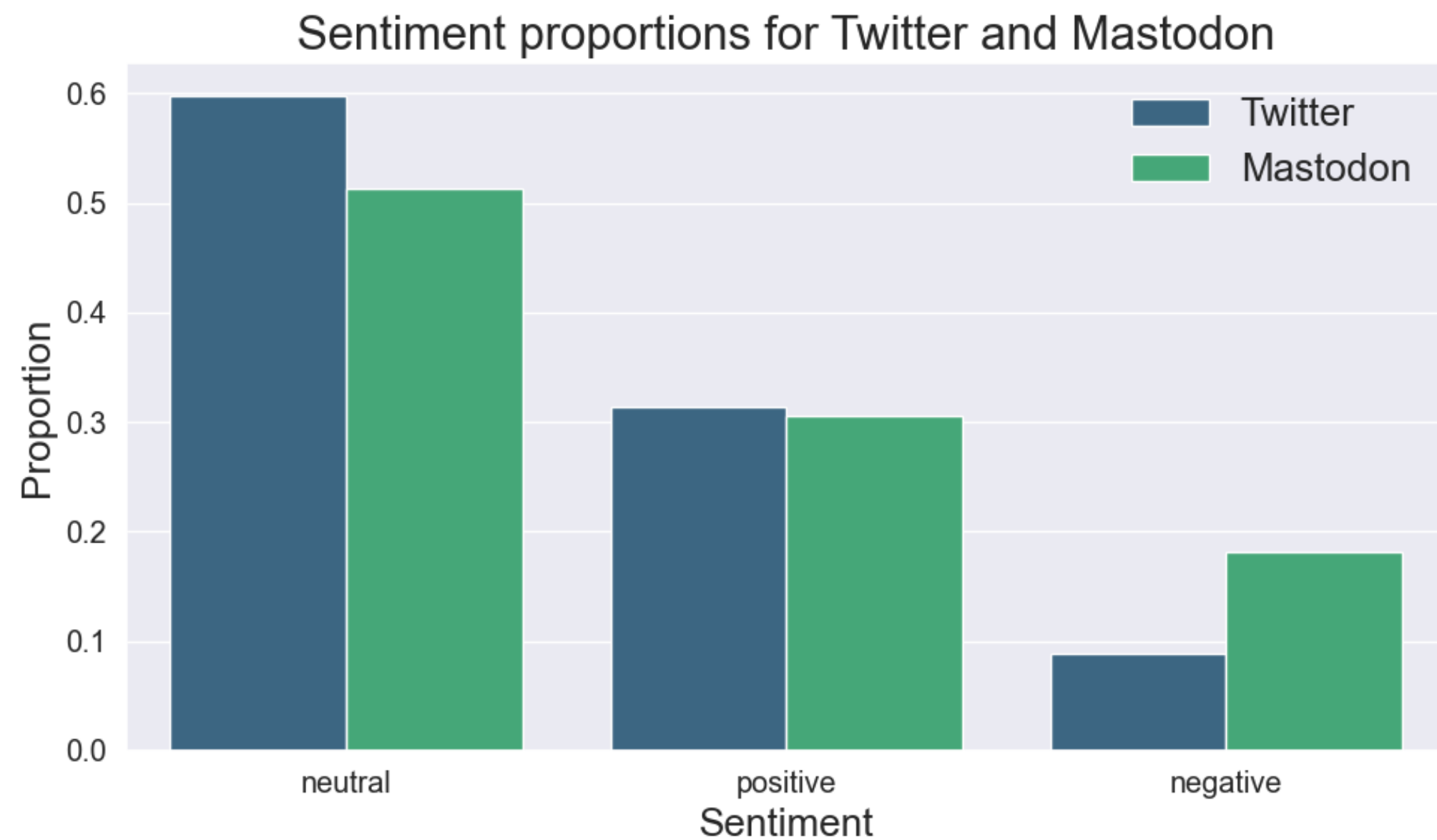
**roBERTa**

Robustly Optimized Biding-tional  
Encoder Representations  
Transformers Approach





# RESULTS



**Twitter** proportions:

neutral: 59%

positive 31%

negative: 9%

**Mastodon** proportions:

neutral 51%

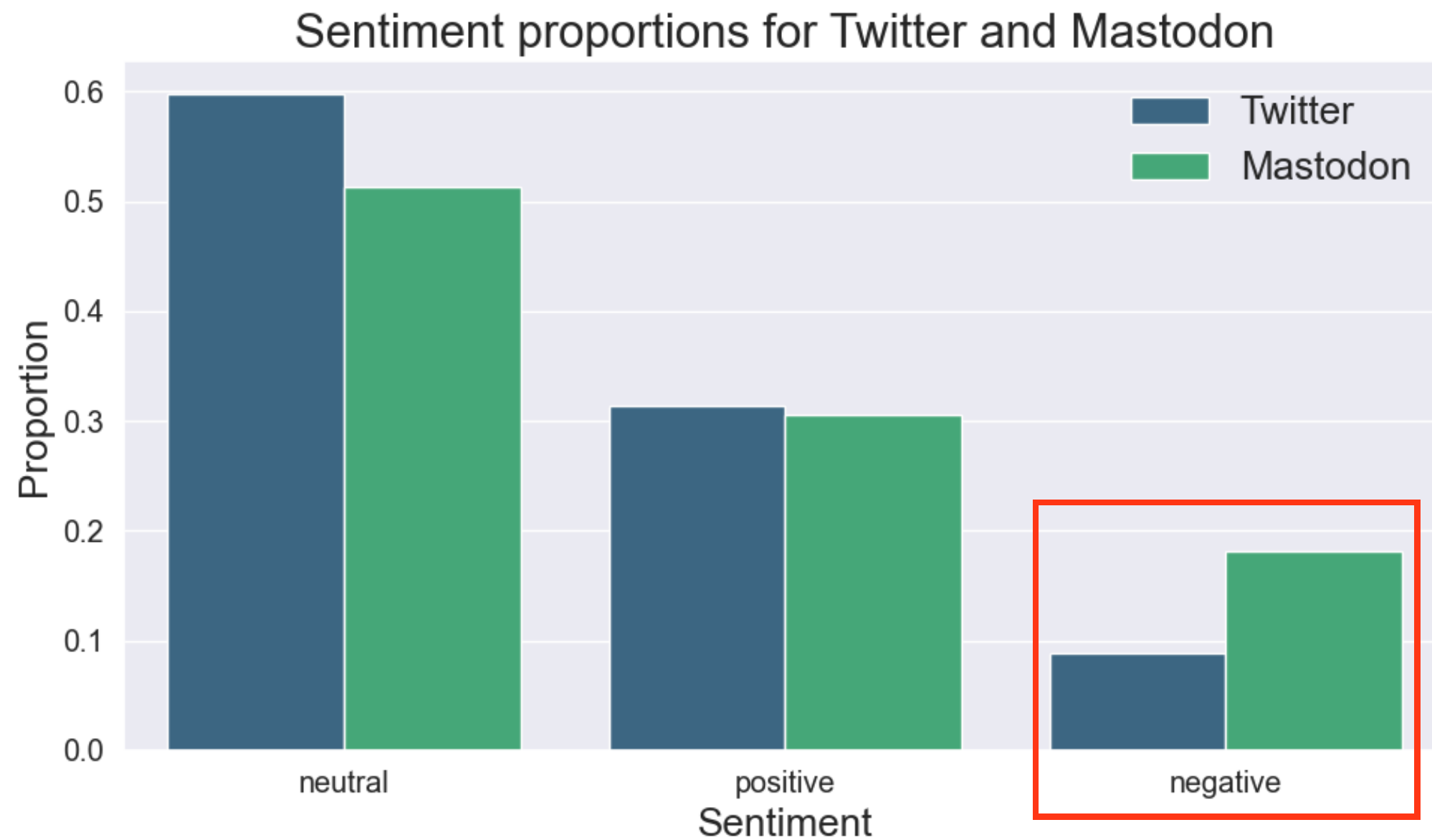
positive 31%

negative 18%

**Chi Square Test**

$p = 2.6978498793808697e-202$

# RESULTS



**Twitter** proportions:

neutral: 59%

positive 31%

negative: 9%

**Mastodon** proportions:

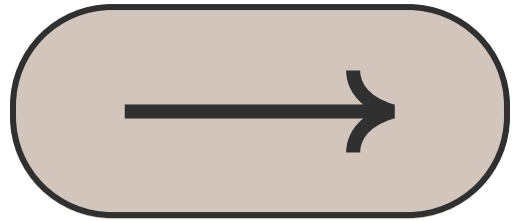
neutral 51%

positive 31%

negative 18%

**Chi Square Test**

$p = 2.6978498793808697e-202$



# Qualitative Analysis

Exploring usability in Mastodon vs. Twitter using  
**Grounded Theory** (Glaser and Strauss, 1967)

- Interviews with hosts
- Documentation of instance's code of conduct / Mastodon / Twitter
- User experiences (Blog posts, Reddit)
- Research publications
- News coverage

## Mastodon <> Twitter

### **Which content do users see?**

How is this perceived by users?

--> polarizing content

--> bias / echo chambers

### **Which content do users not see?**

How is this perceived by users?

--> harmful content

# Which content do users see?

## Mastodon

USERS ONLY SEE CONTENT FROM OTHER USERS ON THEIR INSTANCE

USERS CAN FAVOURITE OR BOOST (=RETWEET) A TOOT BUT WITHOUT ADDING A COMMENT

USERS CAN CONFIGURE THEIR FEED (LOCAL VS. GLOBAL)

FEED IS IN CHRONOLOGICAL ORDER

## Twitter

USERS LARGELY SEE RECOMMENDED CONTENT (FOR YOU PAGE)

USERS CAN LIKE, RETWEET AND QUOTE TWEET WHICH BOOSTS THE POPULARITY OF CONTENT

USERS' FEED IS THEIR FOR YOU PAGE BY DEFAULT

FEED IS DETERMINED BY POPULARITY

# How is this perceived by users?

## Mastodon



### **POLARIZING / SENSATIONAL CONTENT:**

ANTIVIRALITY ATTRACTS SERIOUS POSTERS AND  
DETERS FALSE INFORMATION



.....  
"BORING" AND IT'S DIFFICULT TO GRASP  
WHAT IS RELEVANT AT THE MOMENT



### **BIAS / ECHO CHAMBERS:**

USER IS RESPONSIBLE FOR THEIR EXPOSURE



.....  
SMALLER INSTANCES ARE PRONE TO THIS

## Twitter



### **POLARIZING / SENSATIONAL CONTENT:**

USERS SEE CURRENTLY "RELEVANT" CONTENT FIRST



.....  
USER FEED ARE FILLED WITH UNWANTED CONTENT



### **BIAS / ECHO CHAMBERS:**

RECOMMENDED CONTENT: SHOWN TO EXPOSE USERS TO OPPOSING  
VIEWPOINTS



.....  
CONTENT IS DETERMINED BY RECOMMENDER ALGORITHMS



# Which content do users **not** see?

## Mastodon

USERS DON'T SEE CONTENT FROM OTHER INSTANCES OR ADS

USERS CAN BLOCK OR MUTE OTHER TOOTS, USERS OR INSTANCES

### **CONTENT MODERATION:**

- MASTODON WIDE COVENANT AND ENCOURAGEMENT OF CONTENT WARNING FLAGS
- BUT: EACH INSTANCE HAS ITS OWN CODE OF CONDUCT
- USERS REPORT INCIDENTS DIRECTLY TO THEIR INSTANCE HOST
- CASE-BY-CASE MODERATION IS MOST COMMON

## Twitter

CONTENT IS NOT RESTRICTED: USERS ALSO SEE CONTENT FROM ACCOUNTS THEY DO NOT FOLLOW AND ADS

USERS CAN BLOCK OR MUTE TWEETS AND USERS

### **CONTENT MODERATION:**

- TWITTER WIDE CONTENT MODERATION GUIDELINES (TWITTER RULES)
- USERS CAN REPORT INCIDENTS
- LEVELS OF ENFORCEMENT FROM LIMITING TWEET VISIBILITY TO ACCOUNT SUSPENSION
- AUTOMATED ALGORITHMS + CONTENT MODERATORS

# How is this perceived by users?

## HARMFUL CONTENT:

- INSTANCES ARE CROWDFUNDED WHICH CREATES ECONOMY OF ACCOUNTABILITY
- GENERALLY QUITE STRICT MODERATION
- CONTENT WARNING FLAGS ARE WELL ESTABLISHED

.....

- USERS BANNED FROM ONE INSTANCE CAN JUST MOVE TO ANOTHER
- INSTANCES WITH LITTLE CENSORSHIP DO EXIST
- BUT ARE USUALLY BLOCKED BY MAINSTREAM INSTANCES



## HARMFUL CONTENT:

PROACTIVE REMOVAL OF HARMFUL CONTENT

.....

- USERS ARE DISSATISFIED WITH CONTENT MODERATION AS THEY ENCOUNTER HATE SPEECH, FAKE NEWS AND POLARIZING CONTENT
- LAWMAKERS ARE ALSO DISSATISFIED



**Thank you!**