Exploring Behavioral Tendencies on Social Media:
A Perspective Through Claim Check-Worthiness

2024 International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Zeyu Zhang, Zhengyuan Zhu, Haiqi Zhang, Chengkai Li

Check-worthiness

- Refers to whether a statement contains **factual claims** that the general public would be interested in verifying as true or false

Low Check-worthiness

Relatively low Check-worthiness

High Check-worthiness

Uses of CW

Many works have been focusing on factual claims

- Check-worthy claims detection

- Automated fact-checking

- Rumor detection

- Misinformation diffusion analysis

New Perspective of CW

None of them has linked individuals' behavioral patterns on social media with factual claims

- Different preferences toward factual claims?

- Whether "Birds of a feather flock together" work when considering preferences toward factual claims?

Research Questions

- Do individuals exhibit different behavioral tendencies toward check-worthiness?

- What could be the reasons for the different tendencies toward check-worthiness?

- Do individuals maintain similar levels of check-worthiness between the tweets they post and those they favor?

- Do individuals tend to follow others who exhibit similar tendencies toward check-worthiness?

Types of Tweets

- *Original-tweet* – a tweet that is initially created by a Twitter user

- *Retweet* – a tweet that is reposted by a Twitter user from an original-tweet

- A *Liked-tweet* of X – a tweet liked by user X

Datasets

Dataset	Tweets Count	Users Count	Collection Date	Domain	Description
RSU	40,405,150	11,173	02/2023	Random	Random users and their followees. Each user has up to 100 followees and 3,200 tweets.
HUM	341,285	498	01/2024	Literature, Art, Religion, Philosophy	Users selected from Twitter lists including book authors, artists, Christians, and philosophers. Each user has up to 1,000 tweets.
POL	8,153,745	3,784	05/2023	Politics	Users selected from Twitter lists including congresspeople, senators, governors, cabinet

Analytic Methods

- **ClaimBuster used to compute check-worthiness score for tweets**
  - A deep learning model trained by corpus of presidential debates

  - Output a check-worthiness score (ranged from 0 to 1) to the given text, the higher the more check-worthy the text is

  - API: https://idir.uta.edu/claimbuster/api/

  - Got 88% accuracy on a random sample of 400 manually annotated tweets

```
┌─────────┐         ╱╲                  ┌──────────────────┐
│  Tweet  │───────▶╱ClaimBuster╲───────▶│ Check-worthiness │
└─────────┘        ╲           ╱        │      score       │
                    ╲╱                  └──────────────────┘
```

Analytic Methods
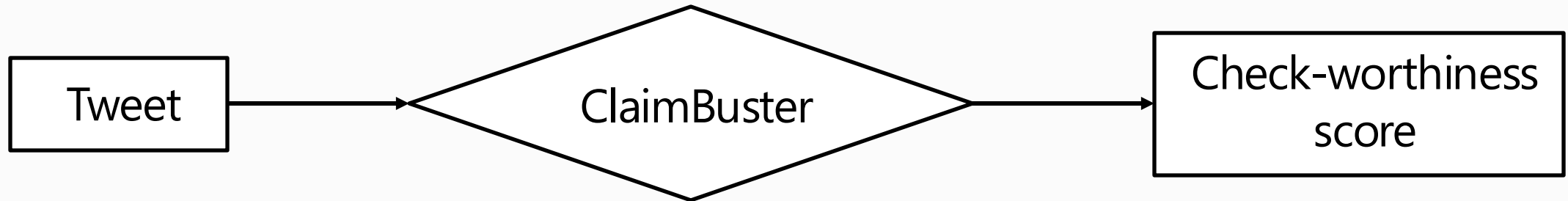
- **Correlation analysis to discover correlation between check-worthiness and various user features**
  - Pearson correlation

- **Hypothesis testing to validate our presumption about individuals' behavioral tendencies toward check-worthiness**
  - Use Brunner Munzel test (BM test)and Kolmogorov-Smirnov test (KS test) because Check-worthiness scores are not normally distributed

Shapiro-Wilk Test of Normality on Check-worthiness Distributions

| H_0 ($\alpha = 0.05$)	Accept	Reject
The check-worthiness scores of a user's original-tweets are normally distributed	37	11136
The check-worthiness scores of a user's retweets are normally distributed	259	10914
The check-worthiness scores of a user's liked-tweets are normally distributed	58	11115

Experiment – Q1

Do individuals exhibit different behavioral tendencies toward check-worthiness?

- **Individual check-worthiness** – using the median check-worthiness of a user's tweets to represent their behavioral tendency toward check-worthiness

- Observe the histogram showing the distribution of individual check-worthiness scores for all users in the RSU dataset

Result – Q1

- Although individual CW mostly concentrates between 0.3 and 0.4, there are people who exhibit a particular tendency toward higher or lower check-worthiness
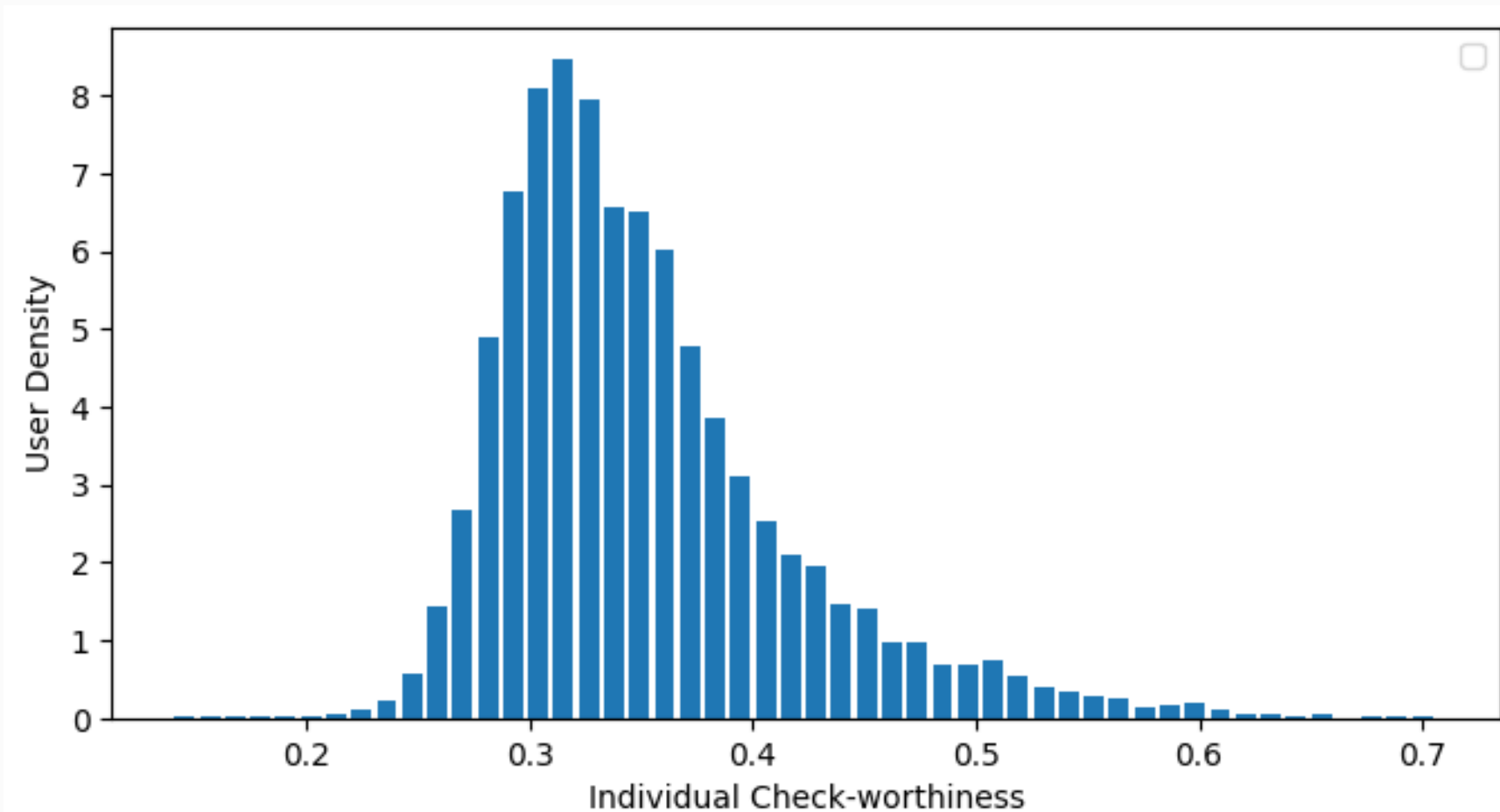
Experiment – Q2

What could be the reasons for the different behavioral tendencies toward check-worthiness? Some features related to popularity and activity level?

- User's Follower Count

- User's Followee Count

- User's Favorite Count

- User's Post Count

- User's Listed Count

- User's Media Count

Correlation Analysis – Q2
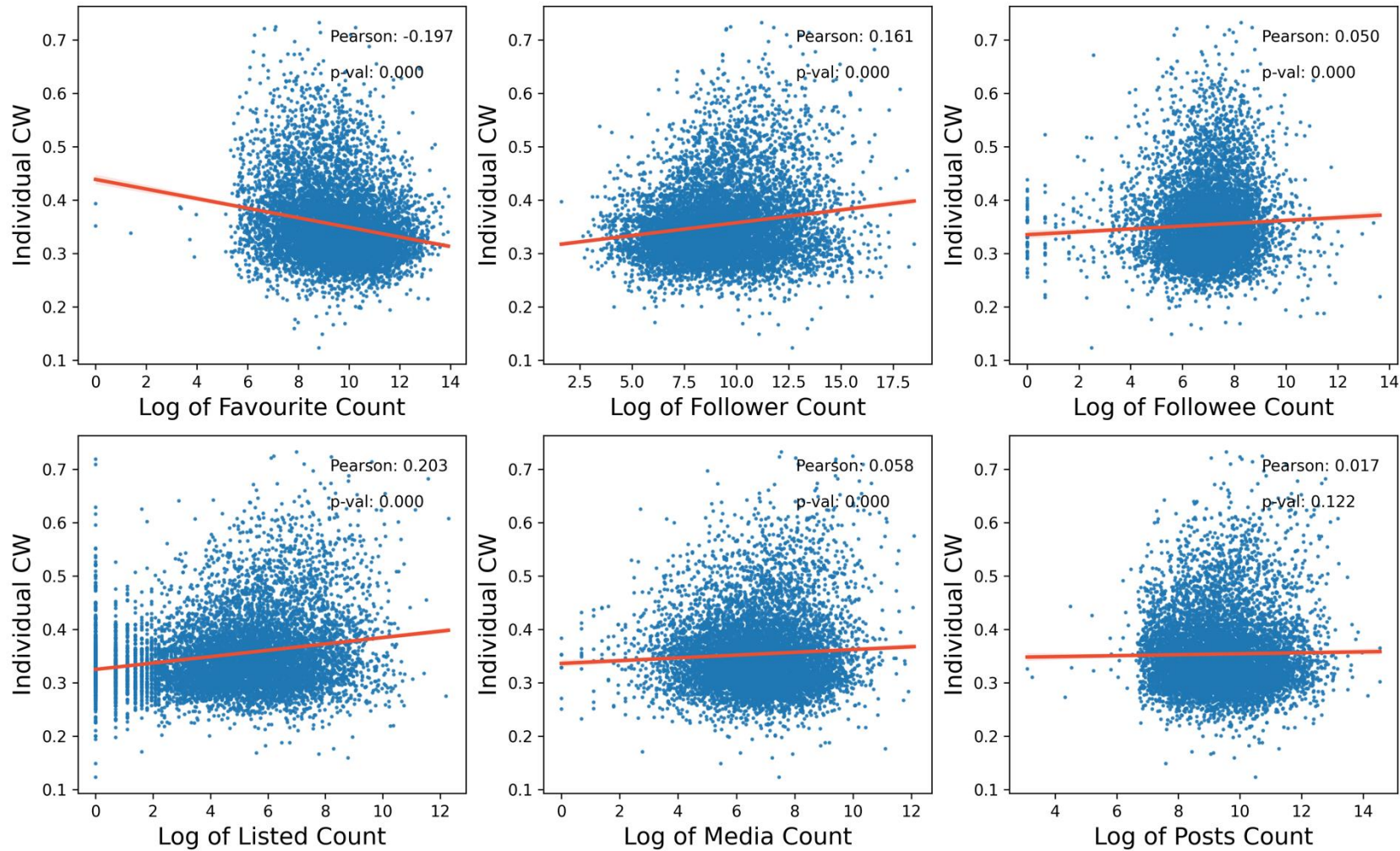
Interpretation – Q2

- All of the features related to popularity and activity level show weak correlations to user's individual CW

- We cannot conclude that those features are indicators of individuals' behavioral tendencies toward check-worthiness
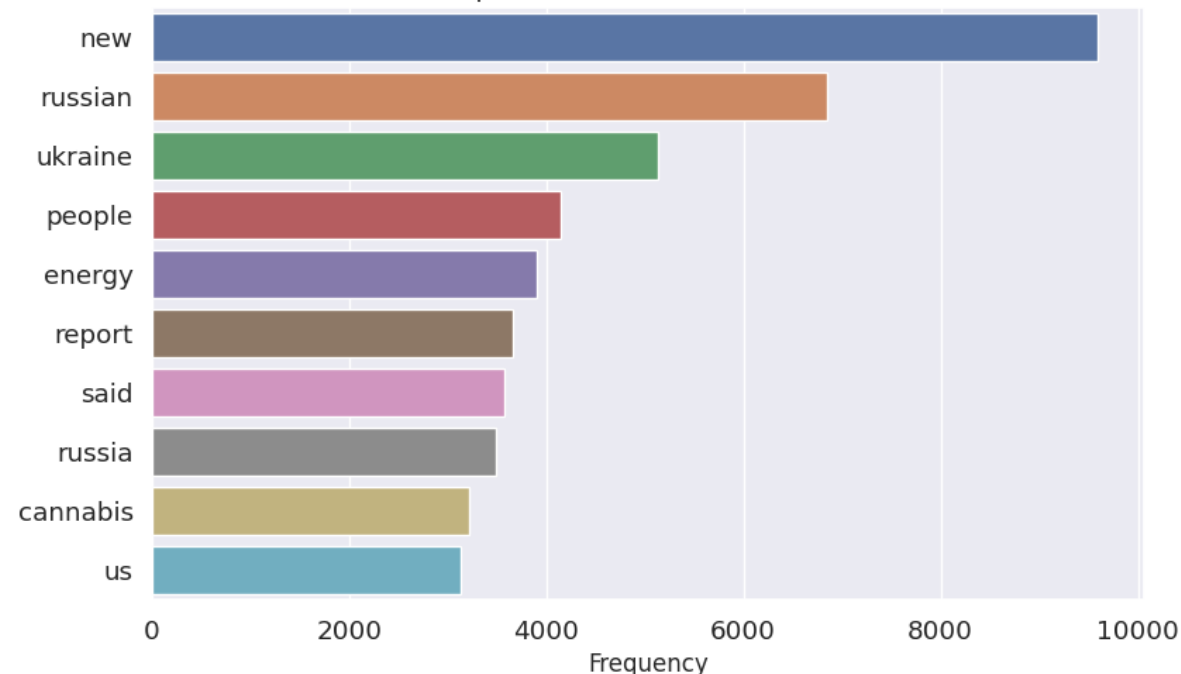
Experiment – Q2

- What about implicit features such as occupation and background?

- Word frequency analysis

  - U_0 – users with individual check-worthiness < 0.25 (146 users)

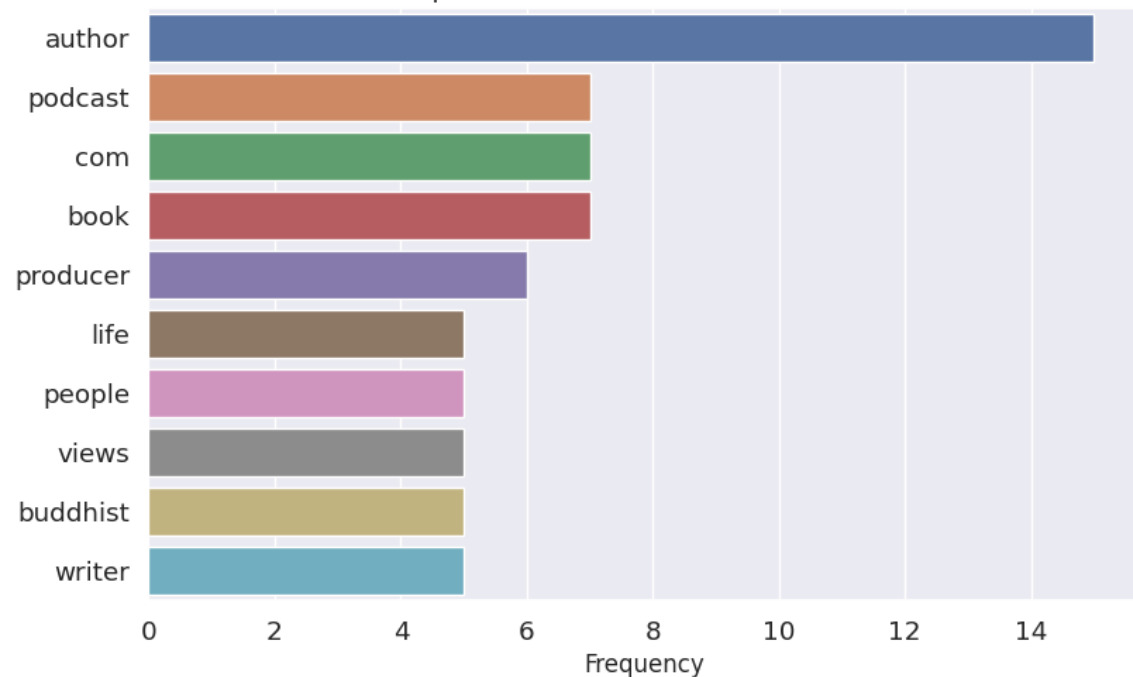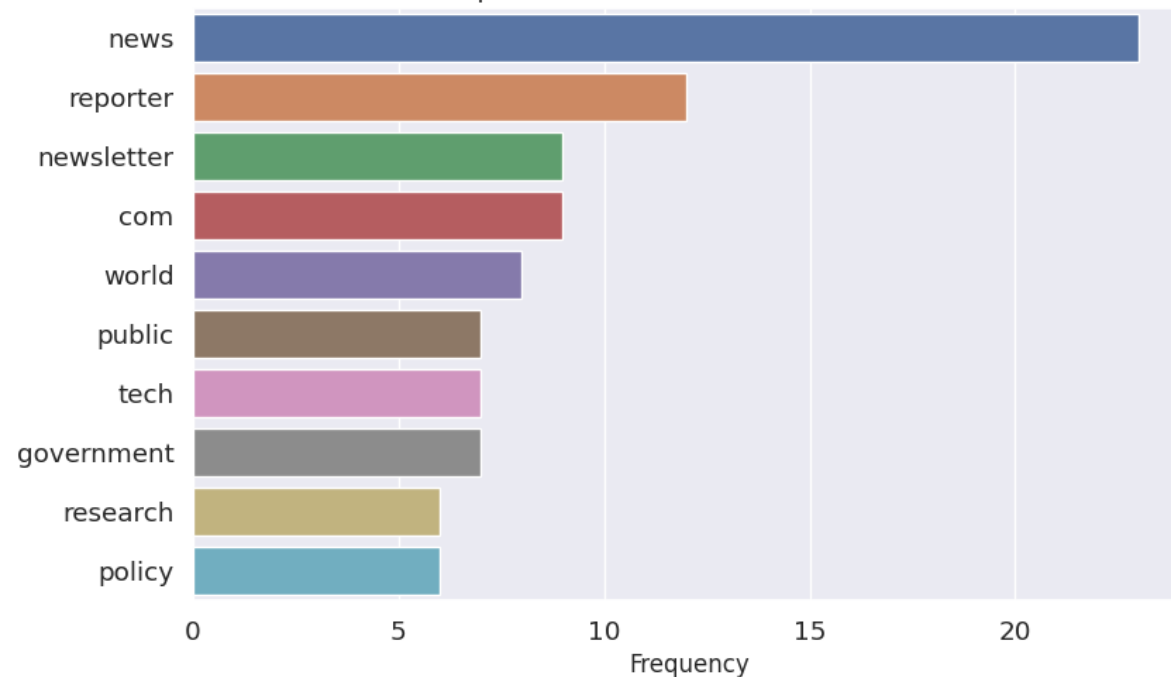 - U_1 – users with individual check-worthiness > 0.55 (169 users)

Top Words in Tweets from U0
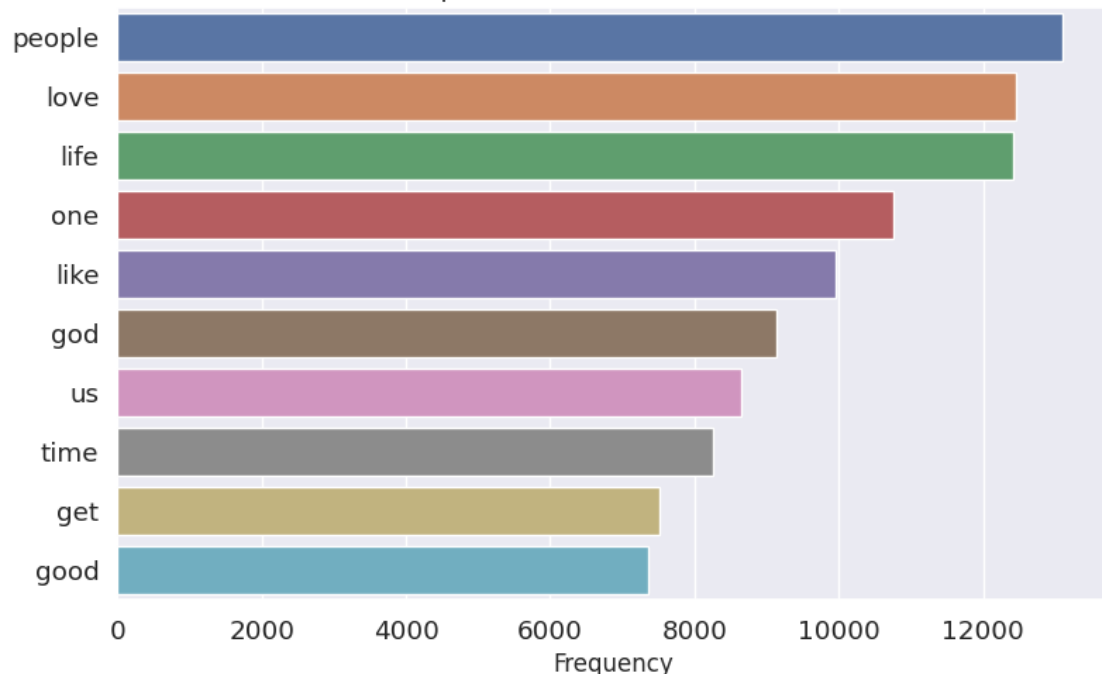
Top Words in Tweets from U1

Top Words in User Profiles from U0

Top Words in User Profiles from U1

16

Interpretation – Q2

- Top frequent words in tweets from U_0 are general and irrelevant to specific people/events/affairs (e.g., love, life, like, god, good)

- Top frequent words in tweets from U_1 are more concrete and highly related to trending topics/events (e.g., russian, ukraine, cannabis)

- Top frequent words in user profile descriptions from U_0 are more related to literature/art, life/entertainment, and religion

- Top frequent words in user profile descriptions from U_1 are more related to journalism, politics, and technology.

- **Individuals' professions, backgrounds, and interests are potentially related to their behavioral tendencies toward check-worthiness？**

Experiment – Q2

- To confirm this conjecture, we randomly selected 100 users from U_0 and U_1 respectively, and then annotated each user account based on their backgrounds and interests.

	Top-ranked Backgrounds/Interests in U_0 and U_1		
U_0's BGs	U_1's BGs	U_0's Interests	U_1's Interests
unknown (36)	media (33)	ideology (40)	politics (25)
writer (19)	reporter (13)	daily life (39)	general news (14)
influencer (12)	research (9)	religion (7)	public good (12)
pastor (3)	politician (6)	entertainment (3)	tech&science (12)
speaker (3)	analyst (5)	photography (2	climate (9
singer (2)	journalist (4)	writing (2)	energy (8)
photographer (2)	unknown (4)	general (2)	security (7)
consultant (2)	writer (4)		business (6)
student (2)	advocate (4)		war (3)
teacher (2)	editor (3)		economics (2)

Interpretation – Q2

- **A majority of users in U_0 lack explicit backgrounds, though a considerable portion comprises writers and influencers**
  - Their primary focus lies in sharing their ideologies and daily lives

- **Users in U_1 are prominently associated with the media, with over half of the selected users actively doing media-related jobs. Besides, users with backgrounds in research and politics are also notably present**
  - The dominant interests within U_1 encompass politics, general news, public interests, and technology/science

- **Align with our previous conjecture**

CW Distributions – Q2

- To further solidify this conclusion, we compared the individual CW distributions of three specific groups of users using the aforementioned datasets – HUM, RSU, and POL
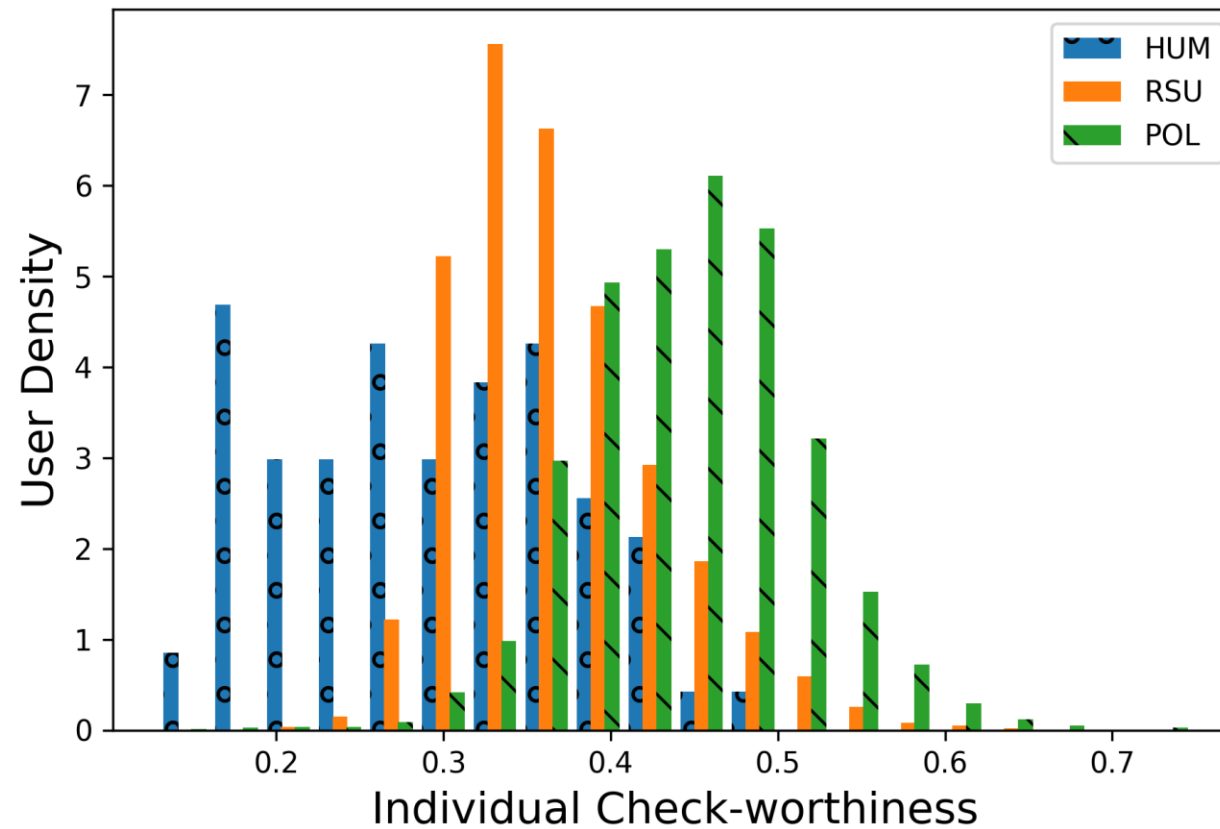
Interpretation – Q2

- The figure shows a left-skewed distribution for the HUM dataset, a right-skewed distribution for the POL dataset, with the RSU in the middle

- Users in the HUM dataset generally possess lower individual CW, whereas those in POL tend to exhibit higher levels of CW

- **Consolidate our conclusion**

Experiment – Q3

Do individuals maintain similar levels of check-worthiness between the tweets they post and those they favor?

- Given a user
    - O – random variable of the check-worthiness of an **original-tweet**
    - R – random variable of the check-worthiness of a **retweet**
    - L – random variable of the check-worthiness of a **liked-tweet**

- Given a dataset
- X – random variable of the check-worthiness of a **random tweet in the dataset**
- P – random variable of the check-worthiness of a **random popular tweet (liked or retweeted by anyone) in the dataset**

Hypothesis Testing – Q3

$$\text{Hyp1} \begin{cases} H_0 : & P(O > R) = P(O < R) \\ H_a : & P(O > R) \neq P(O < R) \end{cases}$$

$$\text{Hyp2} \begin{cases} H_0 : & P(O > L) = P(O < L) \\ H_a : & P(O > L) \neq P(O < L) \end{cases}$$

$$\text{Hyp3} \begin{cases} H_0 : & P(O > P) = P(O < P) \\ H_a : & P(O > P) \neq P(O < P) \end{cases}$$

$$\text{Hyp4} \begin{cases} H_0 : & P(O > X) = P(O < X) \\ H_a : & P(O > X) \neq P(O < X) \end{cases}$$

Acceptances of Hyp1-4

Test	Hyp1	Hyp2	Hyp3	Hyp4
BM Test	1797/16.1%	2896/25.9%	939/8.4%	1013/9.1%
KS Test	1126/10.1%	1831/16.4%	248/2.2%	284/2.5%

Interpretation – Q3

- **The accept rates of Hyp1-2 are greater than that of Hyp3-4 for all the tests**

    - CW of a user's original-tweets is more likely to be similar to CW of the user's retweets and liked-tweets, in comparison to a random and popular tweet from arbitrary users

Correlation Analysis – Q3

- There exist relatively strong correlations between the median CW scores of a users' original-tweets, retweets, and liked-tweets
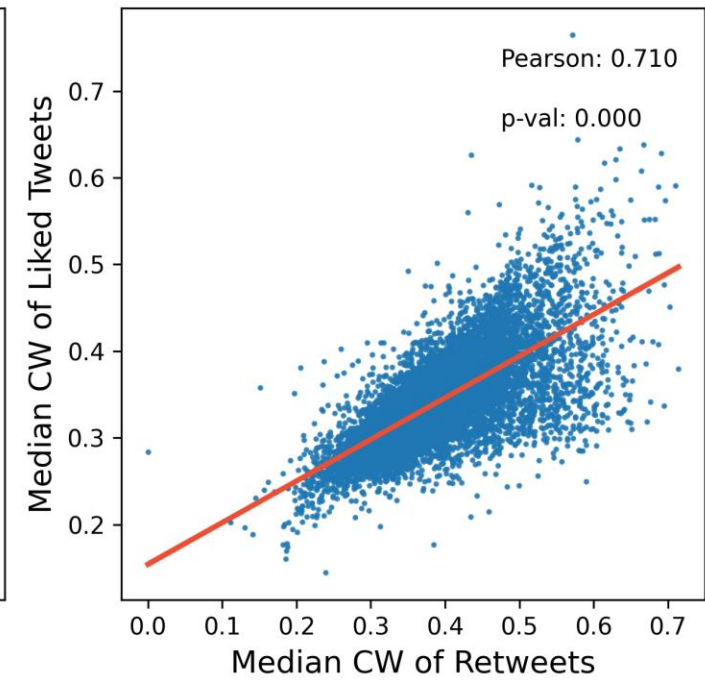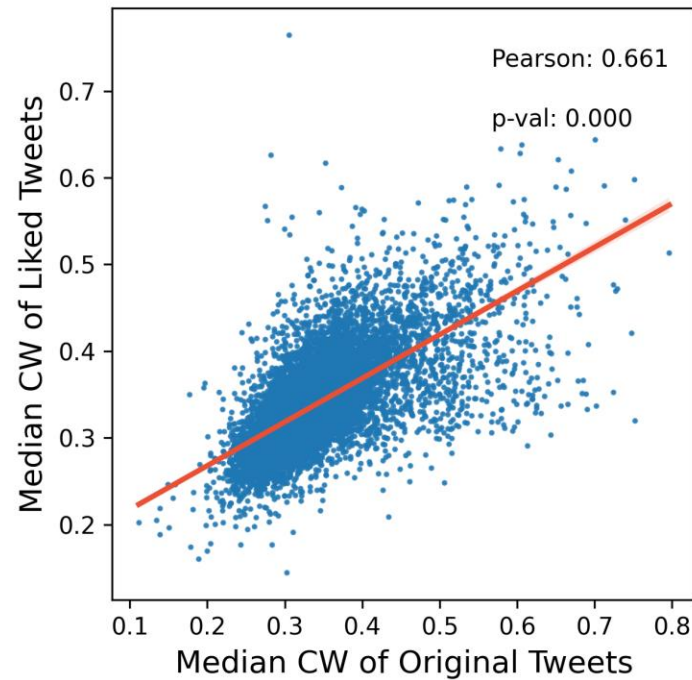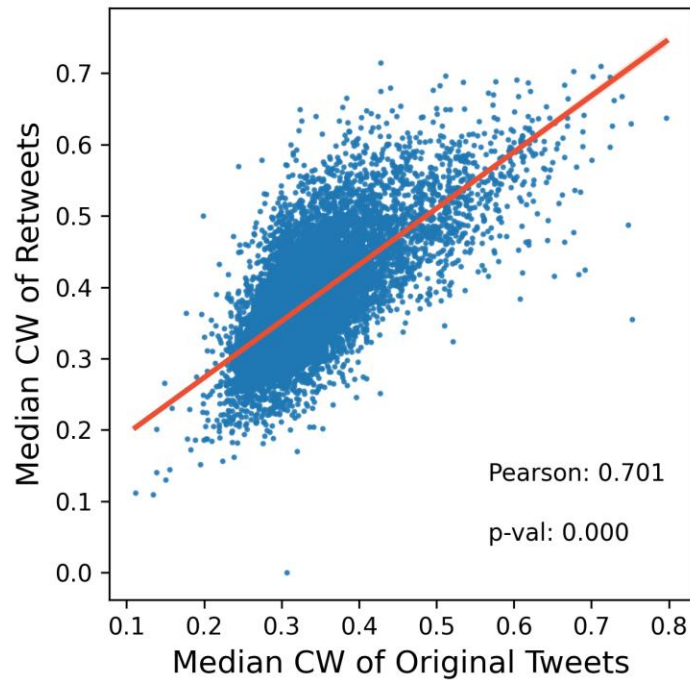
Experiment – Q4

Do individuals tend to follow others who exhibit similar behavioral tendencies toward check-worthiness?

Given a user,

- U – random variable of the check-worthiness of a tweet from **the given user**
- V – random variable of the check-worthiness of a tweet from **one of its followers**
- F – random variable of the check-worthiness of a tweet from **one of its friends (being both follower and followee)**
- X – random variable of the check-worthiness of a tweet from **another random user**

Hypothesis Testing – Q4

$$\text{Hyp5} \begin{cases} H_0: & P(U > V) = P(U < V) \\ H_a: & P(U > V) \neq P(U < V) \end{cases}$$

$$\text{Hyp6} \begin{cases} H_0: & P(U > F) = P(U < F) \\ H_a: & P(U > F) \neq P(U < F) \end{cases}$$

$$\text{Hyp7} \begin{cases} H_0: & P(U > X) = P(U < X) \\ H_a: & P(U > X) \neq P(U < X) \end{cases}$$

Acceptances of Hyp5-7

Test	Hyp5	Hyp6	Hyp7
BM Test	1043/10%	59/16.9%	696/7.6%
KS Test	335/3.2%	50/14.3%	130/1.4%

Interpretation – Q4

- **The accept rates of Hyp5 are greater than that of Hyp7**

  - CW of a user's original-tweets is more likely to be similar to CW of their followees' tweets, in comparison to a random from arbitrary users

  - However, this likelihood is not strong since the acceptance rates of Hyp5 do not exceed Hyp7 by a lot

- **A more substantial result comes from the acceptance rates of Hyp6, which are much higher**

  - Indicates a higher likelihood of check-worthiness similarity between tweets from a pair of users in a two-way following relationship (friendship) than in a one-way following relationship

Correlation Analysis – Q4

- There exists a very weak correlation between the individual CW of followers and followees. The correlation becomes stronger when we compare the individual CW of users with a two-way following relationship (friendship)
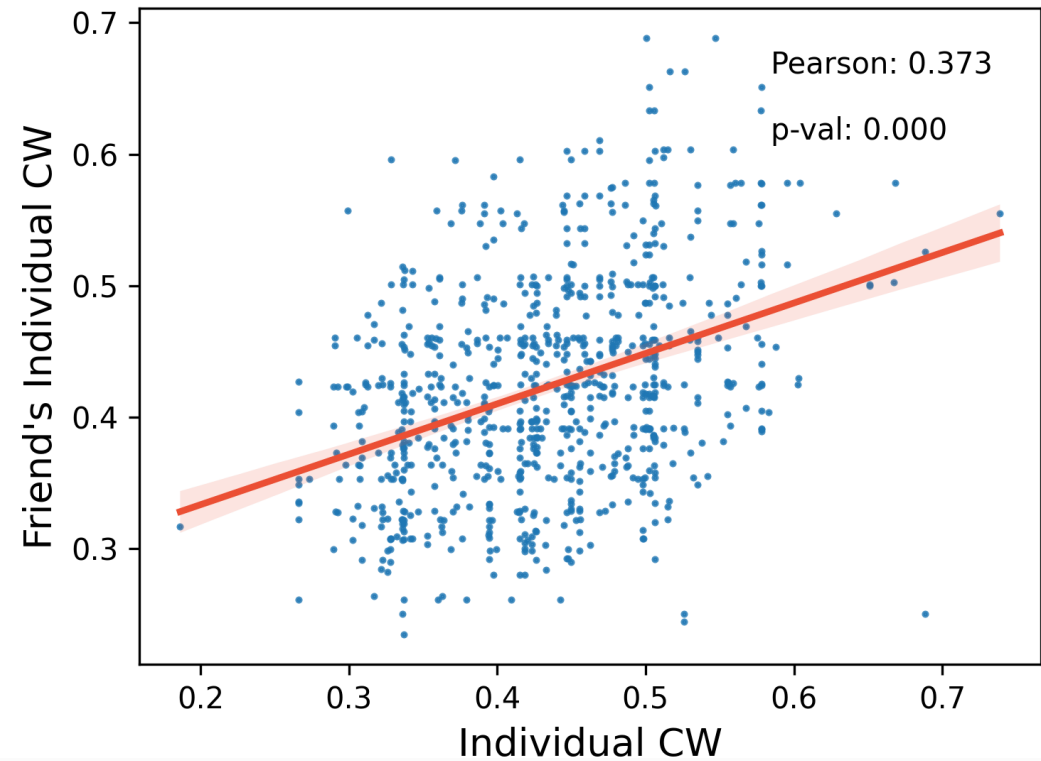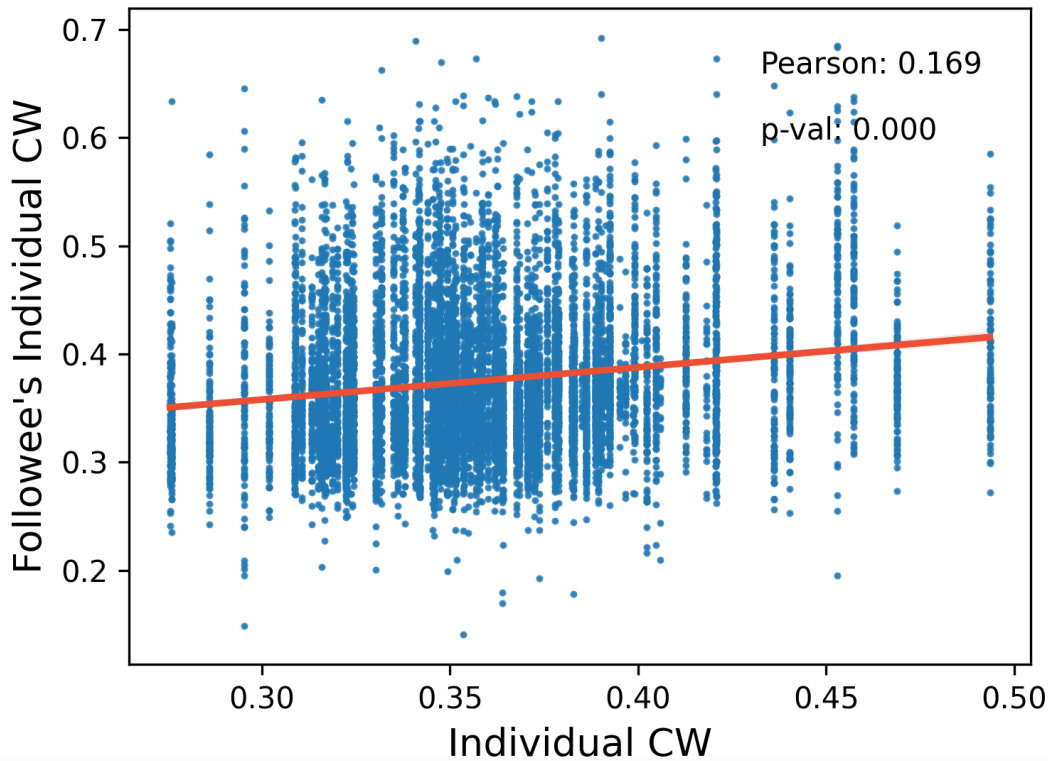
Conclusion

- There exists differences between individuals' behavioral tendencies toward factual claims

- Individuals' backgrounds, occupations, and interests are important factors deciding their tendencies toward check-worthiness

- A strong correlation is discovered between these tendencies and users' posting, sharing, and liking behaviors

- Two-way following relationships (friendships) are likely to reflect shared tendencies towards factual claims