(Dis)information for Hire?

Abstract from old study that should be updated

> Amidst rampant concerns about disinformation, social media influencers (SMIs) can capitalize on their often enormous outreach to spread false claims among their followers. However, despite their sizable potential, the extent to which SMIs sow discord and endorse false narratives is uncharted territory. In this paper, we explore the scale at which SMIs engage with misinformation. We begin by gathering posts from English-speaking influencers with over 500,000 followers on Instagram using CrowdTangle. We then identify instances of disputed content by (i) cross-referencing posts with verified false claims from Politifact, and (ii) manual fact-checking of a random sample of 1000 political posts. This research is pioneering in providing empirical evidence on SMIs' participation in spreading falsehoods. Yet, we find that the concerns are exaggerated, as the involvement of SMIs in propagating false claims is minimal, with only 0.003% of the more than 1.3 million posts analyzed actually supporting statements flagged as disputed by Politifact.

Data

- Instagram Posts from 2000 large entertainment accounts sampled from HypeAuditor, 2020-2023 (`ct_combined_20-23.csv`)
- List of relevant accounts, i.e., excluding entertainment organizations such as sports clubs or movie studios (`account_clusters.csv`)
- List of fake news sourced from Politifact and from Sacha's experiments

Task

The Idea is basically to replicate the procedure proposed by Choi and Ferrarra (2024), probably without the fine-tuning of the LLM.

Codebook

account_clusters

- user_name (str): IG username.
- person (bool): Flag denoting accounts that belong to persons vs accounts that belong to organizations. Here only TRUE, i.e., there are no organizations in the data.
- detailed_description (str): LLM created description based on web search.
- description (str): Description of account from Google Knowledge Graph.
- country (str)
- profile_url (str)
- type_of_account (str): Type of entertainment account, either "Influencers" or "Traditional Celebrities"

ct_combined_20-23

- platform_id (int): Meta ID for account.
- ct_id (int): CrowdTangle ID for account.
- post_created (datetime): Datetime when post was created.
- post_type (str): Type of post (photo, video, album)
- description (str): caption
- post_url (str)
- lang (str): Language of caption (unreliable)
- media_text (str): OCR'rd text from banners in posts.

- overperforming_score (float): CrowdTangle score of performance of post. Basically, a weighed average of the ratios of actual/expected likes and comments.
- update_datetime (datetime): Date when post was last updated in CrowdTangle database. Sometimes identical to post_created - unfortunately there were some glitches in the CrowdTangle data...
- actual_likes (int): Number of likes a post had at update_datetime
- actual_comments (int): Number of comments a post had at update_datetime
- expected_likes (int): CrowdTangle's number of expected likes a post should have at update_datetime
- expected_comments (int): CrowdTangle's number of expected comments a post should have at update_datetime
- full_name (str): Full name of account
- user_name (str): IG user name of account
- profile_pic_url (str): Link to profile picture of IG account. Likely defunct, because IG's CDN-Links are only valid for approx 14 days.
- external_url (str): Public link to IG post.
- verified (bool): Did the IG account have a "blue batch" at update_datetime?
- follower_count (int): Number of followers of account at update_datetime

true_fales_misleading_claims

I am not really sure what all the different variables are. But relevant for us are only two columns:

- f/m/t (str): Is a news headline false, misleading, or true. Spread over distinct sheets.
- headline summary (str): Summary of news headline