

# Fake Deepfakes and Green Screen Scripting

An exploration of #conspiracytiktok content in the era of AI.

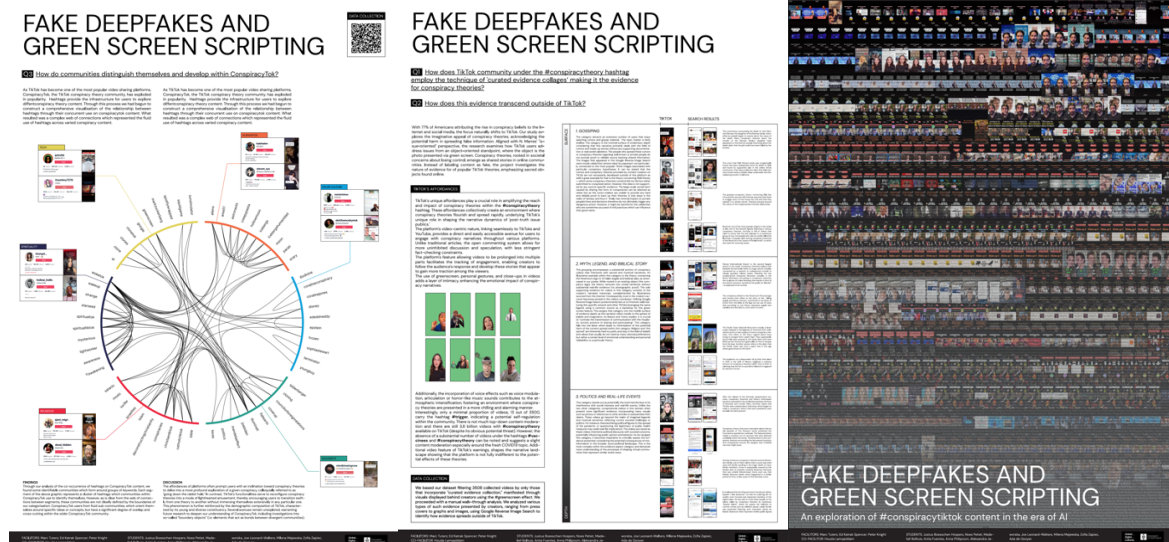
## Team Members

Marc Tuters; Anita Fuentes; Houda Lamqaddam; Aleksandra Jaworska; Joe Leonard-Walters; Milena Majewska; Zofia Zapiec and Ada de Gooyer (with advising from Ed Katrak Spencer & Peter Knight, & thanks for Alice Marwick for sharing her unpublished research)

## Links

Final poster presentations

- [Subproject 1](#)
- [Subproject 3](#)
- [Bonus poster](#)



[Mid-way project presentation google slides](#)

# Summary of Key Findings

The larger project included three smaller sub-projects focusing on 1.) “curated evidence collections” in conspiracy narratives on TikTok (or ConspiracyTok), 2.) using LLMs to categorize the textual content of a large corpus of ConspiracyTok posts, and 3.) using hashtags to categorize ConspiracyTok.

## Subproject 1: TikTok Content Analysis

- **Allure of ConspiracyTok:** ConspiracyTok capitalizes on simplistic power dynamics and the appeal of hidden knowledge with aesthetically appealing visuals prioritized over content reliability, fueling conspiracy theory circulation.
- **Audience Attraction:** Shallow topics on TikTok, particularly those associated with famous figures, tend to attract a large audience, highlighting the influence of celebrity content.

## Subproject 2: Thematic Analysis and Entity Extraction from TikTok Content

- **Beyond Antagonism:** Contrary to expectations, over half of the analyzed TikToks lacked a discernible antagonistic style or "other," challenging assumptions about the prevalence of antagonistic rhetoric in conspiracy-themed content.
- **Limitations of LLMs to categorize ambivalent data:** Many TikToks did not fit the conventional definition of conspiracy theory, complicating thematic categorization and classification via GPT-3.5-turbo
- **ConspiracyTok as ‘Light Entertainment’:** In contrast to other platforms, TikTok presents conspiracies as light entertainment it encourages users to explore alternative realities in a lighthearted and creative manner.

## Subproject 3: Hashtag Network Analysis on ConspiracyTok

- **Hashtag Importance:** Hashtags play a crucial role in distinguishing themes within ConspiracyTok, aiding in the identification of different sub-communities.
- **Concept Sharing:** Sub-communities within ConspiracyTok share concepts and claims despite differentiation, emphasizing the fluid exchange of ideas and themes.
- **Inter-Connectivity Variation:** The level of inter-connectivity among sub-communities varies, highlighting the nuanced nature of ConspiracyTok with distinct clusters forming around specific themes and concepts.
- **Fluid Organization:** ConspiracyTok is characterized by fluid movement across sub-communities, which posing challenges for categorization.

# 1. Introduction

In the midst of the 'Fourth Industrial Revolution' (4IR) and the era of Artificial Intelligence (AI), online conspiracy theories surrounding new technologies and policies have proliferated, offering insights into fears about political power, complicity, and the influence of new media phenomena. TikTok, a prominent platform in contemporary conspiracism, demands critical investigation due to its co-hashtag networks and algorithmic features shaping content on the For You Page (FYP). This introduction aims to seamlessly weave together three interconnected sub-projects, each addressing distinct aspects of ConspiracyTok.

The prevalence of conspiracy theories on TikTok, particularly concerning the music industry, reflects the platform's role in fostering conspiratorial content. The rise of music generative AI and deepfake technology has given rise to conspiracy theories, exemplified by the '#FreeBritney 2.0' phenomenon, linking to broader narratives about the Illuminati's control of the entertainment industry. With a significant portion of the population attributing the rise of conspiracy theories to the Internet and social media, it is imperative to explore how these narratives gain traction. To delve into the grounding and explanation of these theories within TikTok, our study aligns with Marwick et al's concept of "curated evidence collections," (forthcoming) where creators use the platform's features to present theories in the form of memes that travel.

The explosive growth of ConspiracyTok within TikTok, with billions of videos covering diverse topics, necessitates categorization to understand the spread and conjuring of conspiracy theories on the platform. While existing research often focuses on misinformation, particularly around Covid-19, our project shifts attention to the ConspiracyTok community and its limited exploration. Building on Grandinetti and Bruinsma's (2023) claim about emotional elements in conspiracy content, the project centers looks at the affects that connect across ConspiracyTok subcommunities, and attempts to leverage new and emerging methods to map ConspiracyTok.

## 2. Initial Data Sets

This project sought to study so-called ConspiracyTok. TikTok users typically refer to the platform as having different 'zones' such as WarTok or BookTok, which are partially but not entirely defined by that hashtag. Marc Tuters, provided a dataset collected in December 2022 collected using terms hashtags like #ConspiracyTok, "Pizzagate" and "TheGreatReset to manually scrape TikTok using Digital Methods' [Zeeschuimer](#) tool. These were then filtered for only posts containing the world conspiracy which reduced the dataset to 4000. This smaller dataset was then used to provided a substantial sample to gain insights into the nature of ConspiracyTok. The dataset had a temporal bias, due to how TikTok serves data, so that the most recent data included lots of noise. We found that the data from a few months before that date was the 'cleanest'. Thus, while the initial plan was to utilize an extensive dataset of 4,000 TikToks acquired through scraping

between 2021 and 2023, (for subproject 3) a more manageable subset was chosen, resulting in a dataset comprising 150 TikToks collected between October and December 2021, all featuring the keyword "conspiracy" in their textual body. Post-selection, duplicates and unavailable TikToks were eliminated, resulting in a final dataset of 117 posts for subproject 3—a reduced dataset that facilitated hypothesis testing through a comparative analysis of classifications generated by GPT-3.5 (also used in subproject 2) and those derived purely from qualitative analysis (the method used by subproject 1).

## 3. Research Questions

### **Overall Project Focus:**

How do conspiratorial narratives gain attention on TikTok, and what are the sacred objects within the web that provide evidence supporting these conspiracies?

### **Subproject 1:**

- RQ1: How does the TikTok community using the #conspiracytheory hashtag employ the technique of 'curated evidence collages' as evidence for conspiracy theories?
- RQ2: How and with what effect does this evidence transcend beyond TikTok?

### **Subproject 2:**

- Do conspiracy theories on TikTok exhibit antagonistic traits? [RQ 2.1]
- If so, who or what emerges as the primary antagonistic entities within the dataset? [RQ 2.2]
- Can GPT-3.5-turbo generate a precise thematic categorization of the provided conspiracy TikToks? [RQ 2.3]

### **Subproject 3:**

- How do communities distinguish themselves and develop within ConspiracyTok?
- How do #conspiracytok communities arrange themselves around hashtags?
- How rigidly do hashtags create boundaries between ConspiracyTok communities?

## 4. Methodology

### **Subproject 1:**

- Utilized a dataset collected by Marc Tuters, consisting of 3,508 TikTok videos obtained through the Zeeschuimer browser extension under the #conspiracytheory hashtag.

- Filtered the dataset to 566 TikToks that incorporated the greenscreen effect, focusing on "curated evidence collections."
- Conducted a manual walk-through analysis, categorizing videos into three types: celebrity and true-crime gossip, spiritual and mystic narratives, and factual narratives related to political and societal issues.
- Selected 4 videos from each category for a case study, analyzing how curated evidence collections spread outside the platform using Google Reverse Image Search.

### **Subproject 2:**

- Used 4CAT to tokenize the body and hashtags of TikToks, extracting the top 1,000 words for analysis.
- Integrated components suggested by Benedetti and Borra to formulate a prompt for coding the dataset using PromptCompass, choosing GPT-3.5-turbo for its efficiency in entity recognition and classification tasks.
- Assessed accuracy and observed precision in entity recognition but noted instances of entities generated without apparent reasons.
- Examined thematic terms generated by the LLM, finding alignment with hashtags but limited effectiveness due to broadness and lack of descriptive content.
- Identified limitations in discerning an "other" or antagonist within the TikTok corpus due to insufficient information in the textual content of posts.

### **Subproject 3:**

- Initiated a network analysis to understand how sub-communities on ConspiracyTok use hashtags for self-identification.
- Transformed the dataset of hashtags into a network using the 'weighted degree' parameter and applied force-directed specialization for better visualization.
- Employed the Modularity algorithm to identify communities and categorized hashtags into clusters representing sub-communities like Spirituality, Aliens/UFOs, Tech, Government/Institutions, Religious, and Celebrity culture.
- Used Google's Flan t5 chatbot for hashtag categorization and created a visual representation using edge bundling to show connections between top 10 hashtags in each category.

## 5. Findings

### **Subproject 1:**

In our investigation, we closely examined 12 TikToks, evenly distributed across three categories. The first category delved into relatively shallow topics, focusing on gossip about famous figures. The second category explored less-known subjects like legends and myths, often using conspiracy theories as a narrative tool. Lastly, the third category tackled more serious societal issues, including Covid, vaccines, politics, and politicians. Notably, the videos in this category posed a potential risk due to their significant viewership and the potential influence on societal beliefs.

- Shallow topics on TikTok often gather a large audience due to their association with famous figures.
- Conspiracy theories in the second category lacked scientific support and relied heavily on specific legends or stories.
- Serious societal topics on TikTok had the potential to influence public opinion, leading to serious consequences.

### **Subproject 2:**

Our primary goal was to extract entities and thematic terms for analysis, comparing data with the "Identified other" category. However, this experiment faced challenges, leading to manual coding for rectification. Surprisingly, conspiracy-themed TikToks were not predominantly characterized by an antagonistic style. A supplementary test using transcripts aimed to evaluate GPT-3.5-turbo's thematic categorization, uncovering errors and limitations.

- Over half of the analyzed TikToks lacked a discernible villainous "other," challenging the assumption of antagonistic rhetoric.
- GPT-3.5-turbo faced limitations in accurately detecting conspiracy theories, especially in TikToks with visual elements.
- Many TikToks did not align with the definition of conspiracy theory, presenting challenges in thematic categorization and classification.

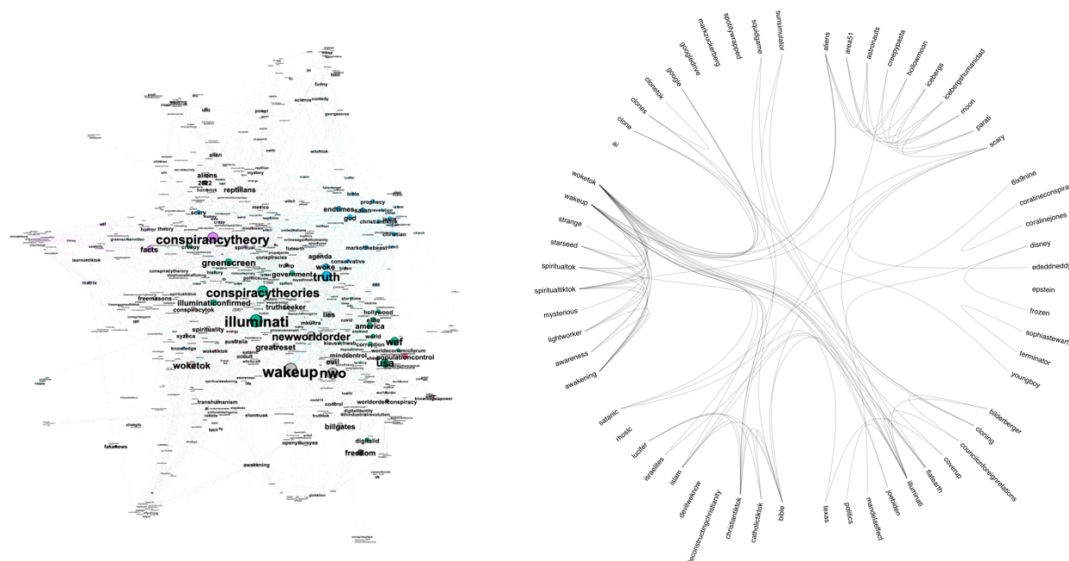


Fig.1 - Network Analysis visualised using Gephi & Fig 2. - Edge Bundling Graph

### Subproject 3:

Our network analysis focused on co-hashtagging on TikTok, revealing the formation of small communities around concurrent hashtag use. Clear clusters emerged, identifying themes within the broader ConspiracyTok community. Notably, a strand of religious conspiracy content was identified through consistent hashtag pairing, leading to the inference of a sub-community around religious conspiracy theories.

- Hashtags play a crucial role in distinguishing themes within ConspiracyTok.
- Sub-communities within ConspiracyTok share concepts and claims despite differentiation.
- The level of inter-connectivity varies among sub-communities, emphasizing the nuanced nature of ConspiracyTok.

## 6. Discussion

### Subproject 1: *Harmful Impact of Conspiracy Theories on Social Matters*

Our research addresses the broader impact of misinformation and conspiracy theories on social life, emphasizing the concept of "harm" in a social context. The project defines harm as having broader social significance beyond personal impacts, particularly in areas such as politics, health, and disasters. Our findings indicate that conspiracy theories on sensitive topics like healthcare and politics have the most harmful impact on social matters. To differentiate the level of harm and evidence required for various conspiracies, we analyzed three recognized categories: private subjects and celebrities, myths and legends, and politics and real-life events.

- Private subjects and celebrities: Rumors and gossip theories lack evidence and primarily exist within the realm of fantasy, causing minimal harm to private individuals but potentially harming celebrities' reputations.
- Myths and legends: Conspiracy theories in this category involve unreal narratives with minimal evidence, linking to beliefs and imagination. Potential harm is vague, as it doesn't significantly interfere with social life and interests.
- Politics and real-life events: This category poses the most potential harm due to its impact on social interests and real-life events. These conspiracy theories often present more significant evidence and can influence public opinion and behavior.

### **Subproject 2: *Antagonism in ConspiracyTok***

Our findings suggest that antagonism is not prevalent across ConspiracyTok as a whole but is associated with specific subsets, particularly conspiracy TikToks centered on governments and institutions. Entities closely linked to the US government, international organizations, and figures from economic elites are identified as antagonists. The level of antagonism varies, with certain entities like Anthony Fauci, NASA, and the FBI featuring prominently as antagonists, while others like the British Royal family and Hollywood are identified to a lesser extent.

- Antagonism is more pronounced in conspiracy TikToks focused on governments and institutions.
- Entities closely linked to the US government, economic elites, and international organizations are frequently portrayed as antagonists.

### **Subproject 3: *Fluidity and Affordances of ConspiracyTok Sub-Communities***

Our analysis indicates that ConspiracyTok organizes itself into fluid sub-communities, representing specific strands of thought within the conspiracy ecosystem. Despite differentiation, all groups interact with the wider conspiracy community, showcasing a high degree of connection. TikTok's unique affordances play a crucial role in amplifying conspiracy theories, with functionalities encouraging a swift transition between theories. The platform's diverse and young user base contributes to the fluid nature of ConspiracyTok, making it challenging to define or categorize.

- Sub-communities within ConspiracyTok are fluid, interacting with the wider conspiracy community.
- TikTok's functionalities encourage a swift transition between conspiracy theories, contributing to the platform's fluid nature.



- The demographic composition of TikTok, characterized by its diversity and youth, influences the nature of ConspiracyTok.

## 7. Conclusion

In navigating the intricate landscape of ConspiracyTok, our collaborative research across three subprojects reveals distinctive facets that contribute to the unique nature of conspiracy theories on TikTok.

### **Subproject 1: *Unraveling the Allure of Conspiracy Theories***

The allure of conspiracy theories, as described by Marwick (2023), capitalizes on simplistic power dynamics and the appeal of hidden knowledge. Our findings emphasize the captivating nature of aesthetically appealing visuals in TikTok videos, prioritized over content reliability, fueling the circulation of conspiracy theories. The creation of online communities supporting these theories thrives on curated evidence collages, amplifying the reach of conspiracy theories. Despite reliable evidence, content creators often reverse scientific confirmation in favor of their own theories, leading to potential misinformation. The impact extends beyond TikTok, infiltrating social contexts and contributing to harmful consequences, particularly in areas related to politics, health, and disasters.

### **Subproject 2: *TikTok as "Light Entertainment" for Conspiracy Theories***

In contrast to other platforms where conspiracy theorizing delves into antagonistic narratives, ConspiracyTok on TikTok presents conspiracies as "light entertainment" (Marwick et al., forthcoming). This approach encourages users to explore alternative realities in a lighthearted and creative manner. The platform's diverse and young user base further contributes to this dynamic, making ConspiracyTok a fluid community where users can transition swiftly between theories without immersing themselves extensively in any particular one. The nature of TikTok's functionalities serves to reconfigure conspiracy theories, fostering a community that prioritizes discussion and knowledge-sharing over anger.

### **Subproject 3: *Fluid Sub-Communities and Unexplored Avenues***

Our analysis indicates that ConspiracyTok organizes into fluid sub-communities representing specific strands of thought within the conspiracy ecosystem. While some groups, like spiritual conspiracy theorists, differentiate themselves using specific hashtags, a significant degree of connection exists among even the most concrete sub-communities. The fluid nature of ConspiracyTok contrasts with platforms like YouTube and Reddit, where users "go down the rabbit hole" into more profound explorations. However, this fluidity poses challenges in defining or categorizing ConspiracyTok due to its diverse and dynamic constituency.

## Common Threads and Future Research

Despite the diversity in our subprojects, common threads emerge in the fluidity, allure, and impact of conspiracy theories on TikTok. Acknowledging limitations, such as dataset constraints, language variations, and the evolving nature of TikTok, prompts us to consider unexplored avenues. Future research should delve deeper into the thematic categorization of ConspiracyTok, community development, and explore "boundary objects" that act as bonds between divergent communities. Our collaborative effort sheds light on the multifaceted dynamics of ConspiracyTok, paving the way for continued exploration and understanding of this unique online phenomenon.

## 8. References

- Baer, N., & Hennefeld, M. (2022). EUROPEAN JOURNAL OF MEDIA STUDIES All the rumors are true: An introduction. *NECSUS*, 11(1), 4–17.  
<https://mediarep.org/server/api/core/bitstreams/2c96f54f-513a-4d63-9ab2-be75e12c79e7/content>
- Bastos, M., & Tuters, M. (2023). Meaningful disinformation: Narrative rituals and affective folktales. *Big Data & Society*, 10(2).  
<https://doi.org/10.1177/20539517231215361>
- Benedetti, A. (2023). Conversing with our machines. Experimental tips on prompting for research with a Large Language Model. DMI Summer School 2023.
- Birchall, C. (2021). The Paranoid Style for Sale: Conspiracy Entrepreneurs, Marketplace Bots, and Surveillance Capitalism. *symploke*, 29(1), 97-121.  
<https://doi.org/10.1353/sym.2021.0006>
- Blackman, L. (2015). The Haunted Life of Data. In G. Langlois, J. Redden, & G. Elmer (Eds.), *Compromised Data: From Social Media to Big Data* (pp. 247-271). London: Bloomsbury Academic.
- Borra, E. (2023). Prompt Compass. <https://promptcompass.digitalmethods.net>
- Born, G., & Haworth, C. (2018). From microsound to vaporwave: Internet-mediated musics, online methods, and genre. *Music and Letters*, 98(4), 601-647.  
doi:10.1093/ml/gcx095
- Cotter, K., DeCook, J. R., Kanthawala, S., & Foyle, K. (2022). In FYP We Trust: The Divine Force of Algorithmic Conspiratorship. *International Journal of Communication*, 16, 1-23. [DOI: 10.1177/14614448231160457](DOI: 10.1177/14614448231160457)

- Enders, A. M., Uscinski, J. E., Seelig, M. I., Klostad, C. A., Wuchty, S., Funchion, J. R., Murthi, M. N., Premaratne, K., & Stoler, J. (2021). The Relationship Between Social Media Use and Beliefs in Conspiracy Theories and Misinformation. *Political Behavior*, 45(2), 1–24. <https://doi.org/10.1007/s11109-021-09734-6>
- Grandinetti, J., & Bruinsma, J. (2023). The Affective Algorithms of Conspiracy TikTok. *Journal of Broadcasting & Electronic Media*, 67(3), 274-293. [DOI: 10.1080/08838151.2022.2140806](DOI: 10.1080/08838151.2022.2140806)
- Hautea, S., Parks, P., Takahashi, B., & Zeng, J. (2021). Showing They Care (Or Don't): Affective Publics and Ambivalent Climate Activism on TikTok. *Social Media + Society*, 7(2). [DOI: 10.1177/20563051211012344](DOI: 10.1177/20563051211012344)
- Mahl, D., Zeng, J., & Schäfer, M. S. (2023). Conceptualizing platformed conspiracism: Analytical framework and empirical case study of BitChute and Gab. *New Media & Society* [Online First].
- Marwick, A., Pippert, C., Furl, K. & Schnabel, E. (forthcoming). Shapeshifters & Starseeds: Identity, Epistemology, and Disinformation in Conspiracy TikTok.
- Peeters, S., & Hagen, S. (2022). The 4CAT capture and analysis toolkit: A modular tool for transparent and traceable social media research. *Computational communication research*, 4(2), 571-589.
- Rosseinsky, K. (2023). #FreeBritney 2.0: Why Britney Spears became a lightning rod for internet conspiracy theories. *The Independent*. <https://www.independent.co.uk/arts-entertainment/britney-spears-teeth-conspiracy-theories-b2434775.html>
- Smith, N., & Southerton, C. (2022). #Freefey and the Pleasures of Conspiracy. *M/C Journal*, 25(1). <https://doi.org/10.5204/mcj.2871>
- Tuters, M. (2023). Everything is literally connected: Studying online conspiracy theories inductively through longitudinal co-hashtag networks. *9th International Ethnography and Qualitative Research Conference > Conspiracy Theories / Conspiracy Practices*. Trento, Italy: University of Trento.
- Valaskivi, K. (2022). Circulation of Conspiracy Theories in the Attention Factory. *Popular Communication*, 20(3), 162-77. <https://doi.org/10.1080/15405702.2022.2045996>
- Zeng, J., & Abidin, C. (2021). '#OkBoomer, time to meet the Zoomers': studying the memefication of intergenerational politics on TikTok. *Information, Communication & Society*, 24(16), 2459-2481. [DOI: 10.1080/1369118X.2021.1961007](DOI: 10.1080/1369118X.2021.1961007)