





# Digital Chameleons: Identity Signal Changes in COVID-19 Discourse on Twitter <sup>\*</sup>

Samantha C. Phillips<sup>1</sup> , Scott Leo Renshaw<sup>1</sup> , Michael Miller Yoder<sup>1</sup> ,  
and Kathleen M. Carley<sup>1</sup> 

IDeaS Center, Software and Societal Systems Department  
Carnegie Mellon University  
5000 Forbes Ave, Pittsburgh, PA 15213  
{samanthp, srenshaw, mamille3, carley}@cs.cmu.edu

**Abstract.** Social media users include identities, a type of signal, in their profile to inform others about themselves and their objectives. These signals emerge and evolve over time due to constant feedback loops and social influence, as well as significant events. In this work, we analyze the frequency and timing of user biography changes of users continuously participating in the COVID-19 discussion on Twitter through 2020 and 2021. Overall, we identify 3,728,076 biography changes in 2020 and 2,479,059 changes in 2021, with spikes around June, July, and October 2020. Yet a higher proportion of users in the COVID-19 communication space changed their bio throughout 2021 than 2020, possibly indicating strategic signaling. Our results have implications for understanding when and why signals used in online self-presentation evolve.

**Keywords:** Identities · Signals · Social media · COVID-19

## 1 Introduction

Social media users present themselves through online social identities, a type of signal, in their profile to inform others about themselves. Users can present themselves through their unique username, what they write in their biographies, and through content in their posts. Without offline signals available, users rely on these modes of self-presentation to curate their online persona and form impressions of other users they interact with or read the posts of.

The identities contained in these self-presentations, often characterized by sociologists as a “front stage” performance tailored to the given context, are

---

<sup>\*</sup> This work was supported in part by the Knight Foundation and the Office of Naval Research grant MURI: Persuasion, Identity, & Morality in Social-Cyber Environments, N00014-21-12749. Additional support was provided by the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Knight Foundation, Office of Naval Research, or the U.S. Government.

not constructed in isolation [3,7]. Signaling theory suggests people use signals to selectively share information with other parties to maximize personal benefits and minimize risks [2]. Risks include social fallout, where that individual may be shunned or removed from a community they identify with [7].

Signals emerge and evolve over time due to constant feedback loops, as well as significant events [2,1]. On social media, shifts in signal use can be captured in real time in a relatively complete sense relative to offline. We aim to shed light on the signal updating process by analyzing the timing and frequency in which people change the signals in their self-presentation on social media. Specifically, we examine user biography (bio) changes in the context of the COVID-19 discussion on Twitter in 2020 and 2021.

COVID-19 was a monumental crisis around the world, sparking controversy about public health measures and looming over events like elections and protests. It affected how people voted and engaged in collective action to further social justice objectives [6,4]. Hence, it is imperative to study how people respond to such events while accounting for the ongoing effects of COVID-19. We restrict our analysis by only considering bio changes of users continuously participating in the COVID-19 discussion on Twitter. This way we are capturing how people who actively post about COVID-19 change their self-presentation in response to ongoing events and signals from other users.

## 2 Methods

### 2.1 Data Collection

Our dataset was collected using a streaming keyword search via Twitter v1 API between January 1, 2020, and December 31, 2022. We collected tweets that contained at least one of the following terms: coronavirus, coronavirus, Wuhan virus, Wuhanvirus, 2019nCoV, NCoV, NCoV2019, covid-19, covid19, covid 19. This results in 468M tweets and 17.2M unique users, where 10.4M users only tweeted one time in our dataset. Figure 1 displays the number of tweets and unique users per month in our dataset.

### 2.2 User bio changes extraction

Since we are not able to directly identify bio change events from the Twitter API, we use the date and time the user tweeted about COVID-19 with each bio (if they tweet with more than one bio) to approximate when users change their bio. For each tweet by each user, we mark if the bio of the user at the time is unchanged from the previous tweet, changed to a new bio, or reverted back to a previously used bio.

## 3 Results

We identified 6,207,135 bio changes between January 2020 and December 2021 by 2,728,024 users participating in discussions about COVID-19 on Twitter. Of

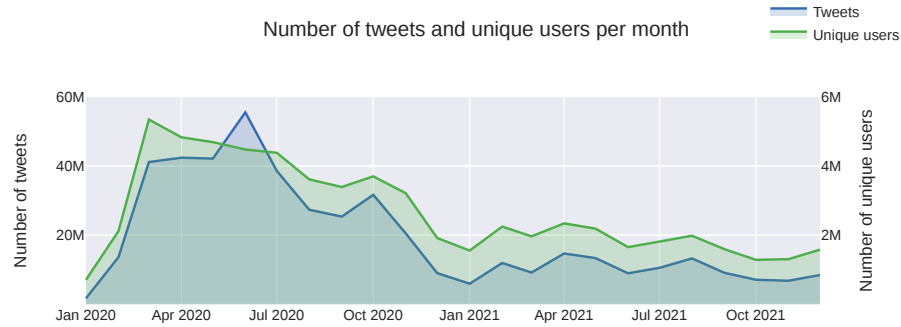


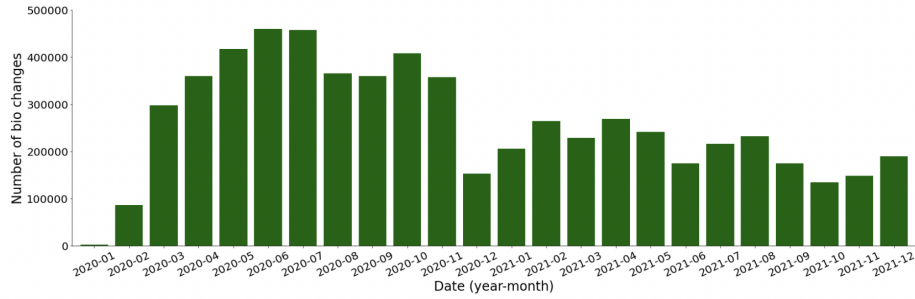
Fig. 1: Number of tweets and unique users over time in our dataset.

these bio changes, 57,677 were reverting to a bio that user previously held. We did not capture any bio changes for 60% of users in our dataset who tweeted more than one time about COVID-19. Figure 2a displays the number of bio changes per month and Figure 2b displays the number of bio changes normalized by the number of unique users participating in the COVID-19 discussion on Twitter per month.

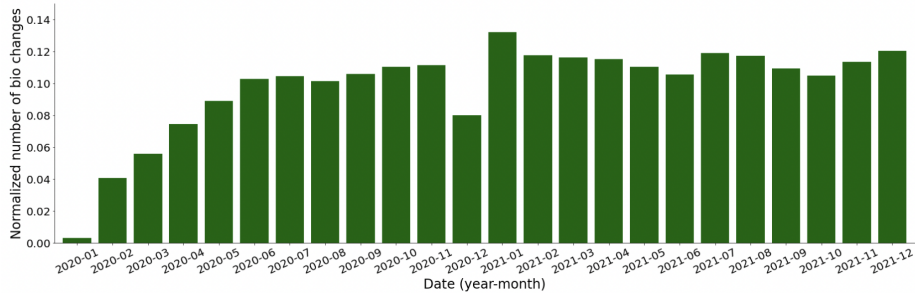
First, consider the counts of bio changes throughout 2020 and 2021 in Figure 2a. The number of bio changes increases January thru June 2020. The peak is 460,434 bio changes in June 2020, followed by 457,589 changes in July 2020. Overall there are more bio changes in 2020 than 2021, aligning with the decrease in overall users discussing COVID-19. In particular, there is a stark difference between the number of bio changes in November 2020 (358,137) and December 2020 (153,277). Unlike 2020, there are less than 300,000 bio changes in every month of 2021.

Now consider the counts of bio changes normalized by the number of users per month in Figure 2b. The highest proportion of users changed their bios in January 2021 (0.13), while December 2020 is a local minimum (0.08). Generally, a higher proportion of users in the COVID-19 communication environment changed their bios in 2021 (avg.: 0.12) than 2020 (avg.: 0.08). This is in contrast to the distribution of raw counts, indicating users continuing to engage with COVID-19 discussions online in 2021 may be strategically signaling within that communication environment.

We also looked at the number of bio changes per user to get a sense of how often users discussing COVID-19 change their bio. The average number of bio changes across 2020-2021 (by users who change their bio at least one time) is 2.4 with a standard deviation of 3.3. Yet 123 users changed their bios more than 100 times during 2020 and 2021 with a maximum of 786 changes by a single user. Changing bios with this frequency may indicate automated activity, especially



(a) Number of bio changes per month.



(b) Number of bio changes normalized by the number of users in the COVID-19 communication environment on Twitter per month.

given most users change their bio 0 or 1 times in this dataset. In fact, 50% of users who change their bio at least one time *only* change their bio one time.

## 4 Discussion

In early 2020, COVID-19 swept the international community and changed the daily lives of billions. People turned to social media and other online forums to express themselves, socialize, and gain information. In this analysis, we examine the evolution of signals used by social media users who turned to Twitter to discuss COVID-19 in 2020 and 2021. Our findings shed light on how people respond to major events and social signals by others in their online social sphere.

The number of bio changes rapidly increased throughout the COVID-19 outbreak from January thru June 2020. The pandemic led to unprecedented public health measures that required effort by citizens. This prompted people to develop new ways to represent their group memberships (or lack thereof) and views, demonstrated by the high number of bio changes during this period. For example, attitudes towards COVID-19 policies became largely intertwined with partisanship in the U.S.<sup>1</sup>

<sup>1</sup> <https://www.science.org/doi/10.1126/sciadv.abd7204>

Although the peak number of users discussing COVID-19 was in March and April 2020, the peak number of bio changes of users in the COVID-19 communication environment was in June and July 2020. This coincides with the Black Lives Matter movement, a sweeping social movement in the U.S. that facilitated protests, some of which violated the social distancing policies and norms of the time<sup>2</sup>.

Examining the number of tweets and unique users discussing COVID-19 over 2020 and 2021, there is a large drop at the end of 2020 as attention shifted away from COVID-19 and towards events like the U.S. Presidential election. The number of tweets about COVID-19 never exceeds 2M after October 2020. Similarly, there is a drop in bio change counts after November 2020.

Notably, a relatively low number of users change their bio in December 2020 despite the growing focus on vaccines following the FDA’s emergency use authorization of the Pfizer-BioNTech COVID-19 vaccine on December 11th. Over time, the salience of COVID-19 waned which makes events like vaccine authorization no longer result in high levels of discussion. If a user stops participating in the COVID-19 communication environment, we do not capture their bio changes.

However, debates around COVID-19 vaccines extend far into 2021. The proportion of users who changed their bio of the users discussing COVID-19 is higher in 2021 than 2020, particularly in January 2021. This could indicate users are strategically leveraging their personas in the COVID-19 communication space. Debates about COVID-19 vaccines contain political, racial, financial, and philosophical components, prompting strong behavioral and attitudinal reactions that also lead to signal updating.

Throughout 2020 and 2021, the local peaks of bio change counts (June 2020, October 2020, February 2021, April 2021, August 2021) are typically followed by a month with more bio changes than the month preceding it. This may be a reverberating effect of feedback between signalers and receivers, as well as the spread of updated signals to other users. The exception is the local peak in August 2021, which is preceded by a month of more bio changes than the following month.

In comparing the number of reversions to previously used bios and changes to new bios each month, we see the relative magnitudes month-to-month are approximately the same. We specifically looked for a disproportionate number of reversions following temporally relevant events like elections when people remove signals designated to support certain candidates. However, that is not represented in our data.

#### 4.1 Limitations and Future Work

We are unable to identify the exact date and time that users change their bio due to data collection limitations. Therefore, we can only capture bio changes of users who continue to participate in the COVID-19 discussion on Twitter. This is highly relevant during the first few months of 2020, when people are learning

<sup>2</sup> <https://www.nytimes.com/article/george-floyd-protests-timeline.html>

about the existence of COVID-19 and posting about it for the first time. We are likely not capturing bio changes during that time.

We also examined the distribution of the number of days between tweets because of the implications for how we extracted the date and time of bio changes. The average number of days between tweets is 49.3 and standard deviation is 108.9 days. The distribution is skewed right with a median number of days between tweets of 5.8 (first quartile is 0.24 and third quartile is 38.0). This indicates the majority of users tweet in bursts, which is enhanced if they are actively interacting with and responding to other users. The users who tweet frequently in a short time frame may not post about COVID-19 again or tweet consistently throughout the two years in our dataset. On the other hand, about a fourth of users talk about COVID-19 sporadically with over a month between their tweets.

Taken together, these results indicate that our results would largely not change if we knew the exactly date/time of the bio change. Given most users tweet about COVID-19 within a month of their previous tweet and our analysis is at a month level, we are likely capturing when bio changes occur within a month of the actual bio change.

The major events of focus in this work are largely U.S.-centered. Although there are more people from the U.S. on Twitter than any other country<sup>3</sup>, there is evidence that people around the world turned to this platform to discuss COVID-19 [5]. Future work should examine *which* political and social signals are being changed each month to strengthen our explanations for bio change frequencies across 2020 and 2021. Furthermore, location indicators in user profiles and posts can assist in explaining bio changes as well.

## References

1. Bliuc, A.M., Betts, J., Vergani, M., Iqbal, M., Dunn, K.: Collective identity changes in far-right online communities: The role of offline intergroup conflict. *New Media & Society* **21**(8), 1770–1786 (2019)
2. Connelly, B.L., Certo, S.T., Ireland, R.D., Reutzel, C.R.: Signaling theory: A review and assessment. *Journal of management* **37**(1), 39–67 (2011)
3. Goffman, E.: *The Presentation of Self in Everyday Life*. Doubleday, Oxford, England (1959)
4. Grant, P.R., Smith, H.J.: Activism in the time of COVID-19. *Group Processes & Intergroup Relations* **24**(2), 297–305 (2021)
5. Ilyas, H., Anwar, A., Yaqub, U., Alzamil, Z., Appelbaum, D.: Analysis and visualization of COVID-19 discourse on Twitter using data science: A case study of the USA, the UK and India. *Global Knowledge, Memory and Communication* **71**(3), 140–154 (2022)
6. Peeples, L., et al.: COVID and the US election: will the rise of mail-in voting affect the result. *Nature* **10** (2020)
7. Smaldino, P.E., Flamson, T.J., McElreath, R.: The evolution of covert signaling. *Scientific reports* **8**(1), 1–10 (2018)

<sup>3</sup> <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>