

POWERED BY TWITTER?

THE TALIBAN'S TAKEOVER OF AFGHANISTAN

Laura Courchesne
Bahar Rasikh
Brian McQuinn
Cody Buntain



JUNE 2022

A joint report by the Centre for Artificial Intelligence, Data, and Conflict,
the Empirical Studies of Conflict Project at Princeton University,
and the New Jersey Institute of Technology

POWERED
BY TWITTER?
THE TALIBAN'S
TAKEOVER OF
AFGHANISTAN
JUNE 2022

The Centre for Artificial Intelligence, Data, and Conflict (CAIDAC) is a research organization dedicated to studying social media's impact on conflict, political violence, and war. We analyze state and non-state actors' evolving social media strategies and their influence on domestic and international perceptions. CAIDAC draws on artificial intelligence research tools to understand how malicious actors weaponize social media and what moderation efforts could curtail dangerous or illegal activities.

CAIDAC is a global network of researchers, practitioners, and humanitarians. The Centre's computer scientists and social scientists integrate their respective knowledge, methods, and worldviews to provide insights into complex problems. Aided by machine learning and artificial intelligence tools, our team strives to find new research approaches through deep integration of their respective disciplines. We collaborate with scholars from the University of Regina, the University of Alberta, the New Jersey Institute of Technology, the University of Maryland, the University of Oxford, and Princeton University. CAIDAC's conflict report series translates academic research into policy briefs that provide analysis of emerging issues.

For further information about CAIDAC or this report, please write to the Centre for Artificial Intelligence, Data, and Conflict at the Politics and International Studies Department, University of Regina, 3737 Wascana Parkway, Regina, Saskatchewan, Canada, S4S 0A2, or visit our website, tracesofconflict.com

Copyright © 2022 by the Centre for Artificial Intelligence, Data, and Conflict
All rights reserved.
Printed in Canada.

Reviewed and copyedited by Tania Inowlocki

Author Acknowledgements

We are immensely thankful for the support, feedback, and comments of Jacob N. Shapiro, Rose B. Huber, Tania Inowlocki, Dorsa Nazemi-Salman, Denilson Barbosa, David Sanger, Shannon Dea, Chris Yost and Eleni Stroulia. Research for this report was supported by funding from the Social Sciences and Humanities Research Council of Canada, the Canadian Department of National Defence Research Initiative, and the University of Regina.



Suggested citation: Courchesne, Laura, Bahar Rasikh, Brian McQuinn, and Cody Buntain. 2022. *Powered by Twitter? The Taliban's Takeover of Afghanistan*. Centre for Artificial Intelligence, Data, and Conflict.

CONTENTS

List of Figures and Tables	5
Executive Summary	6
Introduction	10
I. Taliban Behavior and Engagement on Twitter	12
II. The Taliban's Content Strategy on Twitter	18
III. Twitter's Moderation	21
Twitter's Failed Moderation of Taliban Accounts	21
Twitter's Monetization of the Taliban	22
IV. Policy Implications	25
Notes	25
References	26
Appendix	28

LIST OF FIGURES AND TABLES

Figure 1	Total Tweets by Week from 63 Taliban-Linked Accounts, April 1-September 16, 2021
Figure 2	Number of Taliban Accounts Created per Year, 2011-2021 (n=63)
Figure 3	Total Engagement with Taliban Content per Week, April 1-September 16, 2021
Figure 4	External Links in Taliban Tweets, April 1-September 16, 2021
Figure 5	Mentions (including Retweets) of Taliban-Authored Tweets, April 1-September 16, 2021
Figure 6	Retweeting by Unique Accounts per Day, April 1-September 16, 2021
Figure 7	Daily Use of Keywords in Taliban-Authored Tweets, by Keyword and Language, April 1-September 16, 2021
Figure 8	Daily Engagement with Taliban-Authored Tweets, by Keyword and Language, April 1-September 16, 2021
Table 1	Comparison of Twitter Activity by the Taliban and the 18 Mainstream Afghan News Organizations, April 1-September 16, 2021

EXECUTIVE SUMMARY

On August 15, 2021, a spokesperson of the Islamic Emirate of Afghanistan, the Taliban's self-proclaimed state, declared on Twitter: "With the help of God, and the support of the nation, we are now in control of all parts of the country. We would like to congratulate our nation on this big achievement."¹ After 20 years of conflict with U.S. and NATO coalition forces, no one predicted the speed with which the Taliban would consolidate power and precipitate the collapse of the Afghan government and military.

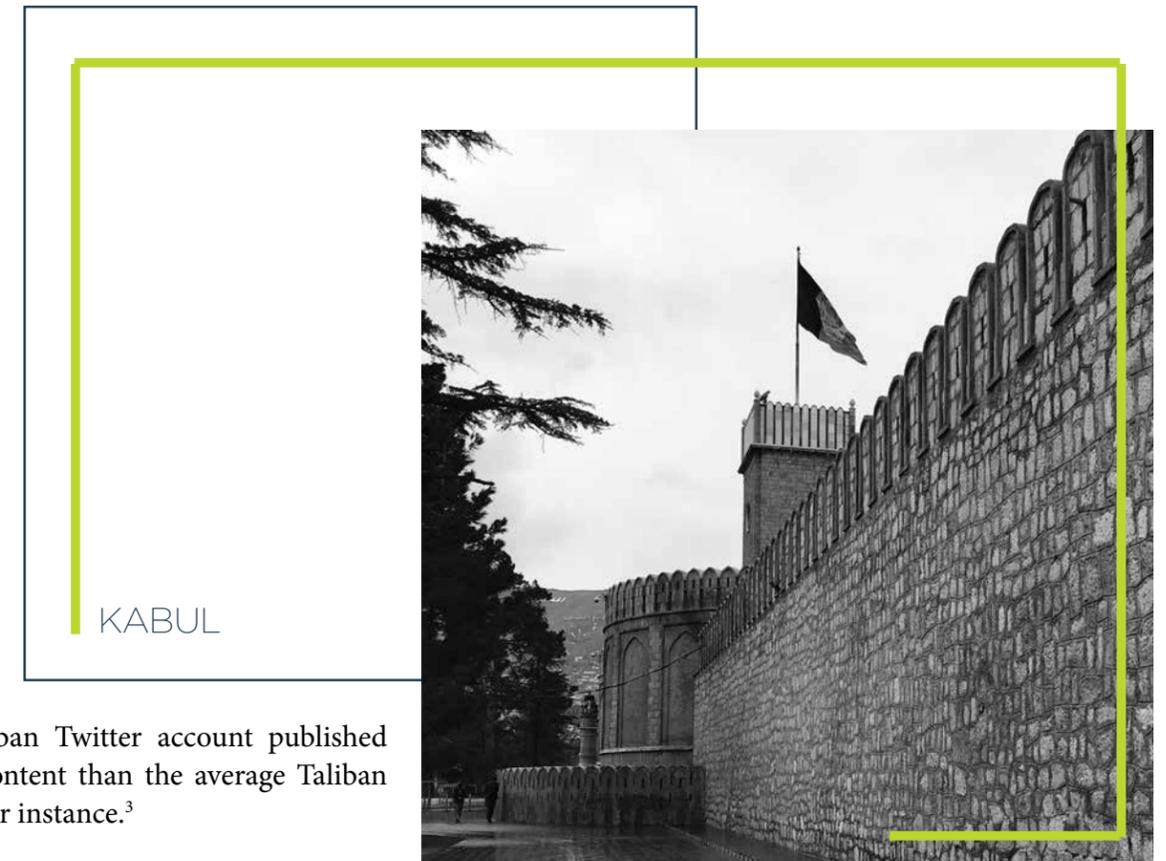
Presenting research conducted by the newly established Centre for Artificial Intelligence, Data, and Conflict (CAIDAC), this report explores social media's central role in the Taliban's takeover of Afghanistan and the strategies used by the group to manipulate international and domestic audiences. It examines the Taliban's social media strategy on Twitter throughout their takeover of Afghanistan, which culminated in the storming of Kabul on August 15, 2021.²

Specifically, we studied the activity of 63 accounts claimed by the Taliban leadership, spokespersons, and avowed members from April 1 to September 16, 2021. These accounts had more than 2 million followers on Twitter in September 2021. As of May 8, 2022, Taliban content reaches more than 3.3 million accounts.

We also examined the broader Taliban ecosystem, which included more than 126,000 Twitter accounts that either retweeted Taliban content or posted content subsequently shared by the Taliban's core network. We found clear patterns in the group's communication strategies, visual imagery deployed, and the timing and content of social media activity and events on the ground. Our analysis supports the following conclusions about the Taliban's Twitter presence:

- **The Taliban weaponized Twitter to dominate Afghanistan's information environment.** The Taliban tweeted well over 100,000 times between April and mid-September 2021. A supportive social media ecosystem of at least 126,000 Twitter accounts then amplified these messages, retweeting Taliban-authored content nearly one million times. The group was so effective at using Twitter to reach domestic audiences that it generated over four times more engagement on the platform than the content of 18 mainstream Afghan news organizations combined. Evidence also suggests that the Taliban chose Twitter as its primary social media platform during the takeover.

6 ¹ See Image 1 in the Appendix for a screenshot of the tweet dated August 15, 2021.
² For further research on the Taliban's social media activity prior to the 2021 takeover, see Bahar (2020), Bernatis (2014), Drissel (2015), Hussaini and Morris (2020), Johnson, DuPee, and Shaaker (2018), and Mehran et al. (2020). For additional reporting and analysis of the Taliban's use of the internet and social media, see Brooking (2021) and Butler (2021).



The average Taliban Twitter account published 23 times more content than the average Taliban Facebook page, for instance.³

- **Twitter appears to have profited from the Taliban's presence on its platform.** Our research indicates that Twitter placed sponsored ads paid for by U.S. and Canadian companies—including Amazon, Disney, McDonald's, and Canadian Imperial Bank of Commerce—on 30% of the 63 Taliban-branded accounts in this report.⁴ Among these accounts were those of the Taliban's news service, spokespersons, and senior leaders. This study also reveals that Twitter continued to place advertisements on accounts that it had flagged for posting "potentially sensitive content."
- **Twitter's moderation failed.** A mere 49 of more than 126,000 accounts in the Taliban support network show evidence of moderation action by Twitter. The vast majority (83%) of Taliban-associated accounts were created before 2021, long before platforms could claim that their presence was permissible because they represented the governing authority in Afghanistan. These accounts also shared content, including graphic images and videos depicting dead and decomposing bodies, in direct violation of Twitter's stated policies on posting and distributing sensitive content. In addition, three-quarters of the Taliban's content was produced by only 20 accounts, suggesting that relatively straightforward moderation efforts could have greatly curtailed the content generated by the group.
- **The Taliban deployed a consistent repertoire of influence strategies to take over Afghanistan.** In concert with their military operations, the Taliban employed six key influence strategies, each with specific goals, imagery, and narratives, to influence international and domestic audiences. The group occasionally promoted plausible—albeit exaggerated or false—assertions. Such disinformation typically involved claims of premature victory or taking undue credit for incidents, all designed to convince Afghans and the international community that the Taliban takeover of Afghanistan was inevitable.

7 ³ This metric comes from an analysis of 21 Taliban-related public Facebook pages gathered from the CrowdTangle API, where the average Facebook page has about 76 posts, whereas the average Taliban-associated Twitter profile posts 1,783 times.

⁴ See Images 2–7 in the Appendix.

- **Since their takeover of Afghanistan, the Taliban have not only kept their accounts active, but also attracted growing numbers of followers, while Twitter has continued to feature sponsored ads on Taliban accounts whose content violates its own policies.**⁵ Twitter continues to monetize the Taliban's presence on the platform. At this writing, the core network had at least 3.3 million followers and ads remained on the account profiles of 19 prominent Taliban accounts. Among the profiles is that of Anas Haqqani, leader of the Haqqani network, which the U.S. government designated a foreign terrorist organization in 2012. Further, many Taliban posts published during the summer of 2021 in violation of Twitter's sensitive content policy remain available.

These findings demonstrate that social media platforms need a better approach to moderation, one combining country-specific tools and methods with a detailed understanding of armed groups and their strategies to manipulate online information landscapes. Blanket bans of groups such as the Taliban are unlikely to have the desired impact, as they can be quickly circumvented.⁶ Instead, effective monitoring strategies require governments and social media companies to engage in ongoing and evolving monitoring, adapting faster than actors who violate applicable laws and policies. In our conclusion, we outline five recommendations to social media companies for improving moderation efforts.

This report has four main sections. The first reviews high-level quantitative trends in the data set. Section II examines the six prominent influence strategies deployed by the Taliban in the lead-up to the U.S. withdrawal. The third section explores Twitter's failure to respond to banned activity, in clear violation of the company's policies, during the months leading up to the Taliban's takeover of Afghanistan. Section IV provides evidence that Twitter profited from the Taliban's presence on the platform by placing advertising on many of the group's accounts. The Appendix presents screenshot evidence of points raised in the report.



5 See the Appendix for examples of Taliban tweets that violate Twitter's policy. Please note that some images are disturbing.

6 See Chakravorti (2021), Ghaffary (2021), and Khilji (2021).

INTRODUCTION

This report offers a preliminary examination of the Taliban’s online information campaign during the takeover of Afghanistan and its immediate aftermath. The period under review begins two weeks prior to U.S. President Joe Biden’s announcement, on April 14, 2021, that U.S. troops would withdraw from Afghanistan; it ends in mid-September of the same year, one month after the Taliban declared victory.⁷

The research supplements media reports and academic analyses of the Taliban’s use of social media by providing quantitative and qualitative analysis of Taliban-linked accounts on Twitter.⁸ Specifically, we examined a core network of 63 accounts, including those associated with Taliban leaders, spokesmen, and avowed members. These accounts published 112,354 tweets in four languages during the time under review. Many profiles explicitly stated their affiliation with and role within the Taliban in their account bios, often in English.⁹ Others included images of the flag and logo of the Islamic Emirate of Afghanistan in their profile pictures or banners while posting insider videos and messages on behalf of the group.¹⁰ By September 2021, this network had amassed more than two million followers on Twitter. The report’s Appendix provides selected screenshots from these accounts.

Weaponizing Twitter as its primary social media platform, the Taliban pursued a sophisticated information campaign to convince Afghans and the international community that its takeover of Afghanistan was inevitable. The group used four primary languages—Dari, English, Pashto, and Urdu—to spread its messaging to specific domestic and international audiences.

Our analysis identified six distinct Taliban narratives, propagated by the Taliban’s network of Twitter accounts and supported by tailored tweets, videos, photos, links, live streams, and Taliban-created hashtags:

1. projecting the Taliban as a government-in-waiting
2. highlighting Taliban military victories by reporting details of specific battles and repeating false claims of success
3. undermining the legitimacy of the Afghan government at the national and provincial levels
4. amplifying or exaggerating the number of civilians killed by the Afghan government and U.S. forces
5. highlighting the defection of Afghan soldiers and their recruitment successes
6. publicizing the Taliban leadership’s meetings and relationships with foreign governments and the international community.

7 See Wagner et al. (2021).

8 For academic research on the Taliban’s historical use of social media and the internet, see Bahar (2020), Bernatis (2014), Drissel (2015), and Mehran et al. (2020). For reporting on the Taliban’s social media behavior during the takeover, see Atiq (2021), Butler (2021), Frenkel and Decker (2021), and Mozur and ur-Rehman (2021).

9 See Image 8 in the Appendix.

10 See Image 9 in the Appendix.

The accounts we identified were created by the Taliban long before they became the governing authority in Afghanistan. The vast majority (82%) of accounts under review were in use before 2021; indeed, Taliban leadership has used Twitter since 2011.¹¹ The broader Taliban ecosystem, composed of supporters who produced content reshared by the Taliban or who actively amplified their content, comprised more than 126,000 accounts. These supporters retweeted or replied to Taliban content nearly one million times in the period under review, and they liked Taliban tweets more than five million times. As a result, the Taliban dominated the information environment on Twitter in Afghanistan. For example, Taliban content received over four times more engagement than 18 mainstream Afghan news organizations combined.

The evidence in this report shows that the Taliban used Twitter more than any other platform, although its presence extended across several platforms. The Taliban published 23 times more content on Twitter than on their corresponding Facebook pages, for instance. Our focus on Twitter reflects the Taliban’s apparent preference for the platform in the months leading up to the takeover and the overt nature of their campaign and engagement with Twitter audiences.

Facebook, Instagram, and YouTube claimed to be taking action against Taliban accounts and activity on their platforms, but the effectiveness or scope of this moderation effort has not been publicly established.¹² Twitter’s current policy towards the Taliban is to moderate accounts and tweets only if they violate its rules against the glorification of violence, information manipulation, or spam behavior.¹³ As this report demonstrates, however, Twitter failed to enforce the rules around the glorification of violence on Taliban-linked accounts. In addition, Twitter’s overall moderation of the broader Taliban supporter network appears to have been limited: our evidence suggests that only 49 of more than 126,000 Taliban-supportive accounts faced moderation action. What is more, it appears that Taliban accounts qualified for the placement of sponsored tweets for companies like Amazon, Disney, McDonald’s, and Canadian Imperial Bank of Commerce.¹⁴

This report has four main sections. The first outlines high-level quantitative trends in the data. Section II examines the six prominent influence strategies deployed by the Taliban in the lead-up to the U.S. withdrawal. The third section explores Twitter’s failure to respond to banned activity, in clear violation of the company’s policies, during the months leading up to the Taliban’s takeover of Afghanistan. Section IV provides evidence that Twitter profited from the Taliban’s presence on its platform by placing advertising on many of the group’s accounts. The Appendix presents screenshot evidence of points raised in the report.

11 See Hendrix (2021) and Singer and Brooking (2018). For research on the Taliban’s use of Twitter in 2012, see Bernatis (2014).

12 See BBC (2021) and Culliford (2021).

13 See Ghaffary (2021).

14 See Images 2–7 in the Appendix.

TALIBAN BEHAVIOR AND ENGAGEMENT ON TWITTER

Using Twitter’s academic research application programming interface (API), we collected tweets posted on prominent Taliban accounts from April 1 to September 16, 2021, a period that covers four and half months before the Taliban’s capture of Kabul and one month thereafter. While many accounts were active before this period, we chose April 1 as a start date to establish baseline behavior prior to U.S. President Joe Biden’s announcement, on April 14, that U.S. troops would withdraw from Afghanistan before the twentieth anniversary of 9/11.¹⁵ We included tweets in the month after the fall of Kabul to document any initial shifts in online behavior following the Taliban’s return to power. Future reports will analyze changes in the Taliban’s information campaign.

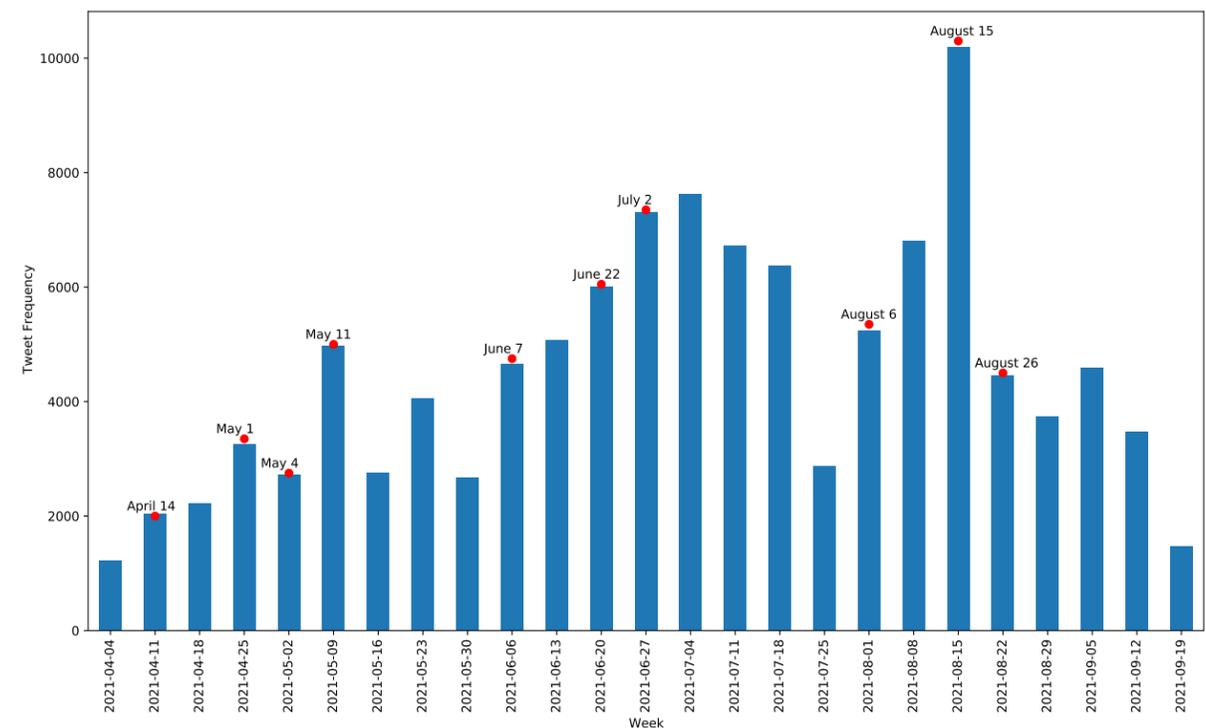
The report studies a network of 63 prominent Taliban Twitter accounts, all of which self-identified as affiliated with the Taliban. In most cases, the account biographies declared users’ positions or affiliations with the group (usually in English) including, for example, the hashtag #taliban in their account descriptions. Accounts also displayed the Taliban logo or flag in profile pictures and banners, or posted tweets that contained messages from or on behalf of the group. These accounts made frequent use of English-language hashtags such as #SuccessesOfTaliban and #WeStandWithTaliban, highlighting the degree to which these accounts operated in the open. A few of the accounts self-described as journalists, although their account bios included a link to the Taliban’s official website, Alemarah.¹⁶

From April 1 to September 16, 2021, the 63 prominent Taliban-linked accounts analyzed in this report produced 112,354 tweets, more than half of which (66,736) were original posts, while the remainder were retweets, replies, or quote tweets. Most of these posts were in Dari, English, Pashto, or Urdu. We also found accounts in the Taliban’s broader amplification network that posted pro-Taliban content in Arabic, Farsi, and Russian. Our analysis reveals a high degree of campaign centralization: the Taliban’s ten most active accounts accounted for nearly half (46.5%) of all tweets, while the top 20 accounts were responsible for more than two-thirds (68.9%) of the Taliban’s content.

The Taliban’s capture of Kabul on August 15, 2021, corresponded with the peak in their Twitter activity. As Figure 1 shows, additional spikes in tweets corresponded to other key dates in the conflict, suggesting close coordination between the group’s social media and military operations. Pivotal dates included:

- April 14: President Biden’s announcement of troop withdrawal
- May 1: a previously established date for total U.S. troop withdrawal
- May 4: the launch of a major Taliban offensive
- May 11: the Taliban’s capture of Nerkh district, outside of Kabul
- June 7: escalation of fighting, with 150 Afghan soldiers killed in 24 hours
- June 22: the Taliban launch of a series of attacks in the north
- July 2: U.S. withdrawal from Bagram Air Base
- August 6: the Taliban’s capture of the first provincial capital
- August 15: the takeover of Kabul
- August 26: the suicide bombing at Kabul Airport by Islamic State Khorasan (ISIS-K).

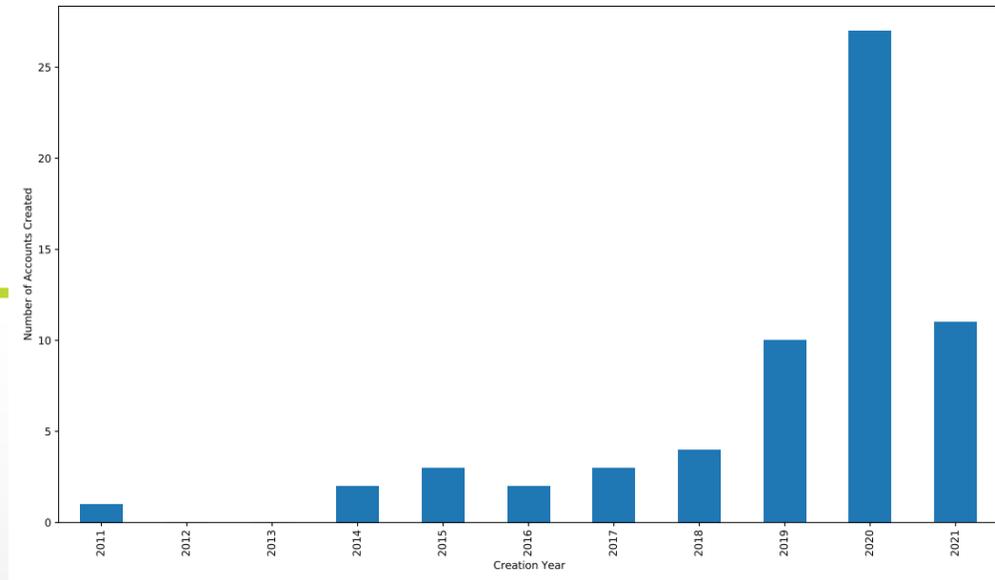
Figure 1 Total Tweets by Week from 63 Taliban-Linked Accounts, April 1–September 16, 2021



15 See Wagner et al. (2021).

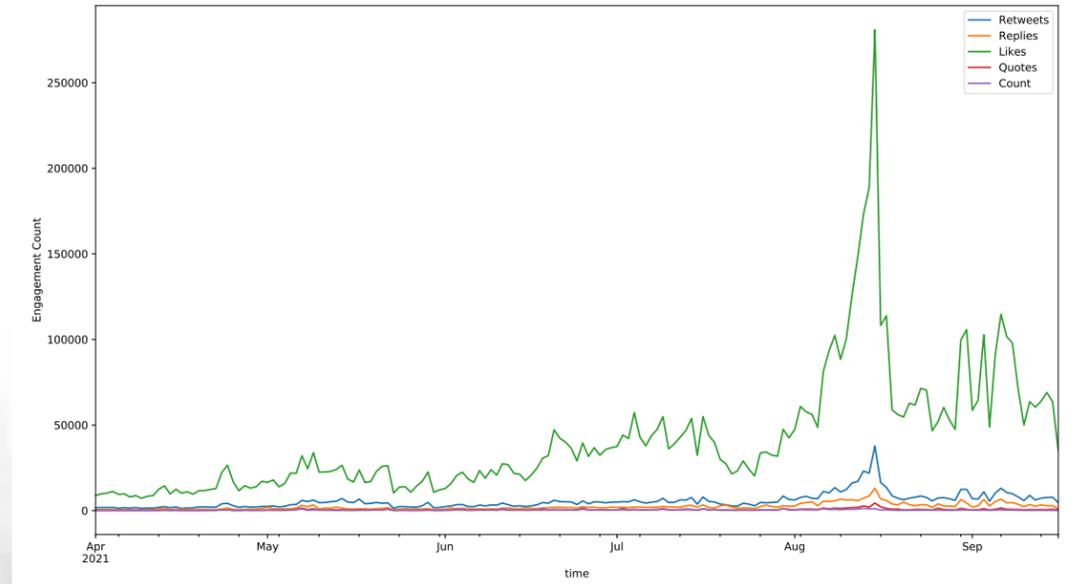
16 The site appeared to go offline in late August 2021. See Timberg (2021).

Figure 2 Number of Taliban Accounts Created per Year, 2011–2021 (n=63)



The Taliban's presence on Twitter predates the takeover by more than a decade. Of the Taliban-linked accounts studied in this report, 83% were created before 2021, long before platforms could claim that the Taliban's presence was permissible because they represented the governing authority in Afghanistan. Figure 2 shows that nearly half (42.9%) of the 63 Taliban-controlled accounts were created in 2020, and that the earliest accounts were created in 2011.

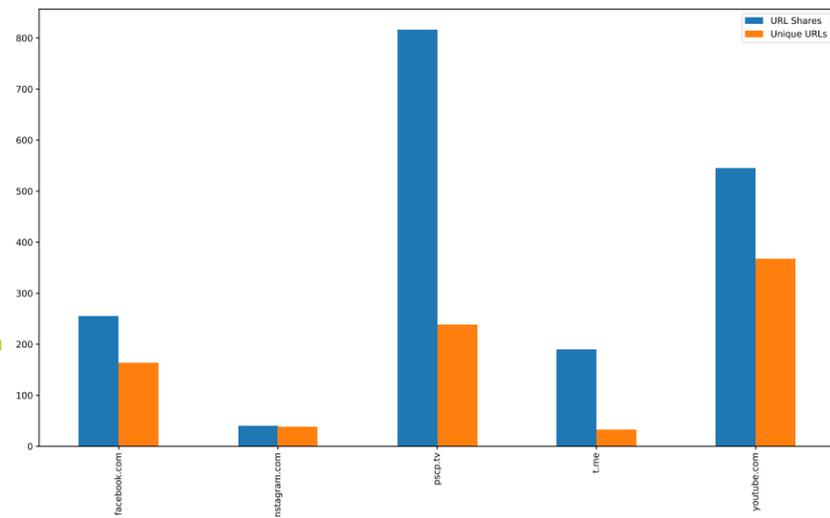
Figure 3 Total Engagement with Taliban Content per Week, April 1–September 16, 2021



To assess the Taliban's influence on the media in Afghanistan, we measured the level of engagement Taliban content received during the period under review. From April 1 to September 16, 2021, Taliban tweets elicited well over 8 million responses: 6,928,223 likes, 940,688 retweets, 387,184 replies, and 93,937 quote tweets. Retweets and likes soared in mid-August, around the time Kabul fell. Figure 3 shows engagement activity involving content authored by the 63 core accounts, including retweets, quote tweets, and replies.



Figure 4 External Links in Taliban Tweets, April 1–September 16, 2021



Note: Blue bars represent the total number of Taliban tweets that included links to external platforms. Orange bars represent the number of unique URLs shared, regardless of how often they were shared.

We also measured how frequently the Taliban used tweets to link to Twitter’s live-streaming feature and other social media platforms, including Facebook, Instagram, Telegram, and YouTube. Only 5.9% of Taliban tweets included external hyperlinks, suggesting that Twitter was the Taliban’s primary social media platform. Of these external links, only 14.6% were directed to other mainstream social media platforms (Facebook, Instagram, Telegram, and YouTube). In other words, less than 1% of the Taliban’s Twitter content involved links to other social media platforms. As shown in Figure 4, the Taliban primarily shared links to Twitter’s live-streaming (10.1% of tweets with links) and YouTube (6.8%). They provided significantly fewer links to Facebook (3.2% of tweets with links), Telegram (2.4%), and Instagram (0.5%). YouTube had the largest volume of unique URLs.

Engagement with the 63 Taliban-controlled accounts saw a steady increase between April and early August 2021, before a massive spike in mid-August, which corresponded to the takeover of Kabul (see Figure 5). Engagement over those four months involved 1,508,701 mentions of Taliban accounts. While not all mentions necessarily represent support, 62% (939,529) of them were retweets, all of which amplified Taliban messaging. Engagement peaked on August 15, with 57,159 mentions of Taliban accounts, 37,753 of which were retweets.

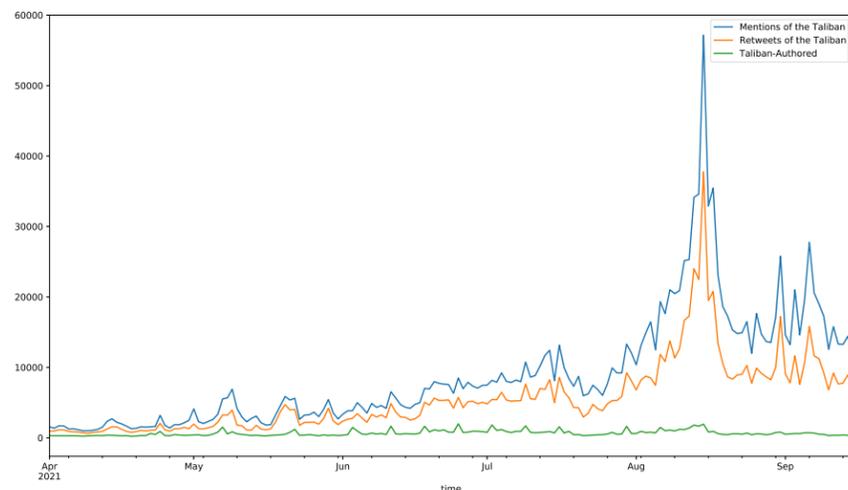


Figure 5 Mentions (including Retweets) of Taliban-Authoring Tweets, April 1–September 16, 2021

Diving further into this retweet behavior, Figure 6 shows how many unique accounts retweeted Taliban messaging every day during the period under review. This line broadly mirrors the mention and retweet patterns in Figure 5. This data shows that May 19, May 20, and May 29 witnessed significant boosts in the Taliban’s retweeting audience, although August 15 saw the most significant influx of new retweeting accounts.

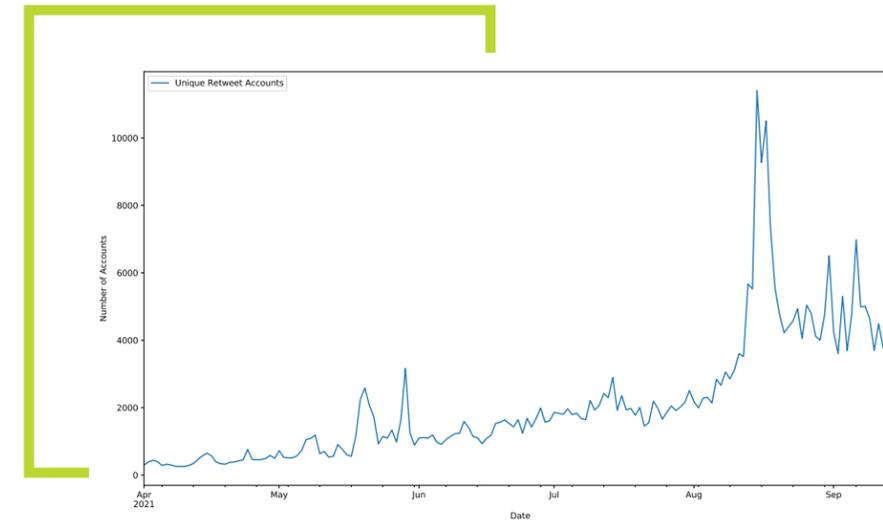


Figure 6 Retweeting by Unique Accounts per Day, April 1–September 16, 2021

During the period under review, the Taliban’s Twitter content received over four times more engagement than that of 18 mainstream Afghan news organizations—including Ariana News, Bakhtar News Agency, Khaama Press, Pajhwok Afghan News, and TOLONews.¹⁷ On average, Afghan media Twitter accounts received fewer than 4 retweets and around 30 likes per tweet; by contrast, the average Taliban account received 14 retweets and 104 likes per tweet. Cumulatively, the 18 news organizations received 1.6 million likes on their content, whereas the Taliban garnered nearly 7 million (see Table 1).

Twitter Activity	Afghan News Accounts	Taliban*
Total tweets	52,054	66,736
Quote tweets	54,148	93,937
Replies	136,392	387,184
Retweets	201,307	940,688
Likes	1,607,290	6,928,223
Mean tweets per account	3,809	1,783
Average retweets per tweet	3.867	14.09
Average likes per tweet	30.877	103.8

Table 1 Comparison of Twitter Activity by the Taliban and 18 Mainstream Afghan News Organizations, April 1–September 16, 2021

Note: * This analysis includes only Taliban-authored tweets posted by the 63 accounts studied in this report, not retweets.

¹⁷ The 18 media organizations analyzed were: Afghan Analysts Network (@AANafgh), Afghan Islamic Press (@aip_news), Afghanistan Sun (@afghanistansun), Afghanistan Times (@AfghanistanTime), Afghan News Now (@AfghanNewsNow), Afghan Voice Agency (@avapress), Ansar Press (@ansarpress1), Ariana News (@ArianaNews_), Azadi Radio (@pazadiradio), Bakhtar News Agency (@bakhtarna), Daily Outlook Afghanistan (@DailyOutlookAfg), Gandhara (@GandharaRFE), Khaama Press (@khaama), Pajhwok Press (@pajhwok), Radio Azadi (@DaRadioAzadi), The Kabul Times (@thekabultimes), TOLONews (@TOLONews), and Wadsam (@wadsamnews).

III. THE TALIBAN'S CONTENT STRATEGY ON TWITTER

Based on our analysis, the following narratives emerged as six distinct strategies in the Taliban's online information campaign in the months leading up to the withdrawal of U.S. troops. For the purposes of this report, a "strategy" includes any regular use of keywords and imagery advancing a specific narrative to a targeted audience. The Appendix presents examples of tweets that correspond to each of the strategies:

1. The Taliban projected themselves as a government-in-waiting (see Images 10 and 11 in the Appendix).
2. The Taliban highlighted military victories (see Images 12 and 13).
3. The Taliban undermined the legitimacy of the Afghan government at the national and provincial levels (see Images 14–16).
4. The Taliban recorded and amplified the mistakes and deaths caused by U.S. and Afghan military forces (see Images 17–20).
5. The Taliban highlighted their recruitment successes and defections from the Afghan military (see Images 21–23).
6. The Taliban drew attention to their relationships with foreign governments and the international community (see Images 24 and 25).

The research also shows how the group integrated disinformation throughout its information operations. While we initially expected more disinformation in Taliban tweets, the group took a more nuanced approach, embedding exaggerations and falsehoods in streams of more plausible information. In July 2021, for example, the Taliban posted an interview with Helmand's Taliban-appointed "mayor." In that interview, he claimed that the group controlled the province, even though they would not be in charge for another month. This interview was part of a series of messages that wove false claims of victory into otherwise accurate announcements, in an effort to influence the morale of Afghan government forces in areas holding out against Taliban attacks. Similarly, in July 2021, the Taliban claimed to have shot down a government helicopter. Even though that assertion was later debunked, it strengthened the narrative of the group's military successes (see Image 13).¹⁸

The bulk of Taliban messaging targeted Afghan civilians, presenting conquered areas as a newly established utopia. Celebratory tweets included video clips, including testimonies from individuals claiming to be from a newly captured city. The point is not that these testimonies were authentic but rather that a significant

portion of the Taliban's influence operations were designed to project the cities they controlled as safe.

The Taliban rarely named their opponents in tweets, opting instead for terminology that roughly translates to "mercenary enemy," "enemy savages," and "occupiers." Some posts, while written in Pashto or Dari, also included the English-language hashtag targeted at criticizing the Afghanistan government, #KabulRegimeCrimes (see Image 20 in the Appendix).

To explore Taliban messaging across different languages, we measured how often three specific keywords appeared in their English, Dari, and Pashto tweets: "victory," "surrender," and "defeat." The analysis used the different languages as proxies for a tweet's target audience. For example, we treated tweets written in English as explicitly targeting international audiences, whereas posts in Dari and Pashto were probably aimed at Afghan citizens. The analysis shows that the Taliban's English-language messaging focused far more on "defeat" and "surrender" than on "victory." The opposite was true for domestic audiences. This discrepancy highlights the extent to which the Taliban tailored their messaging to the intended audience.

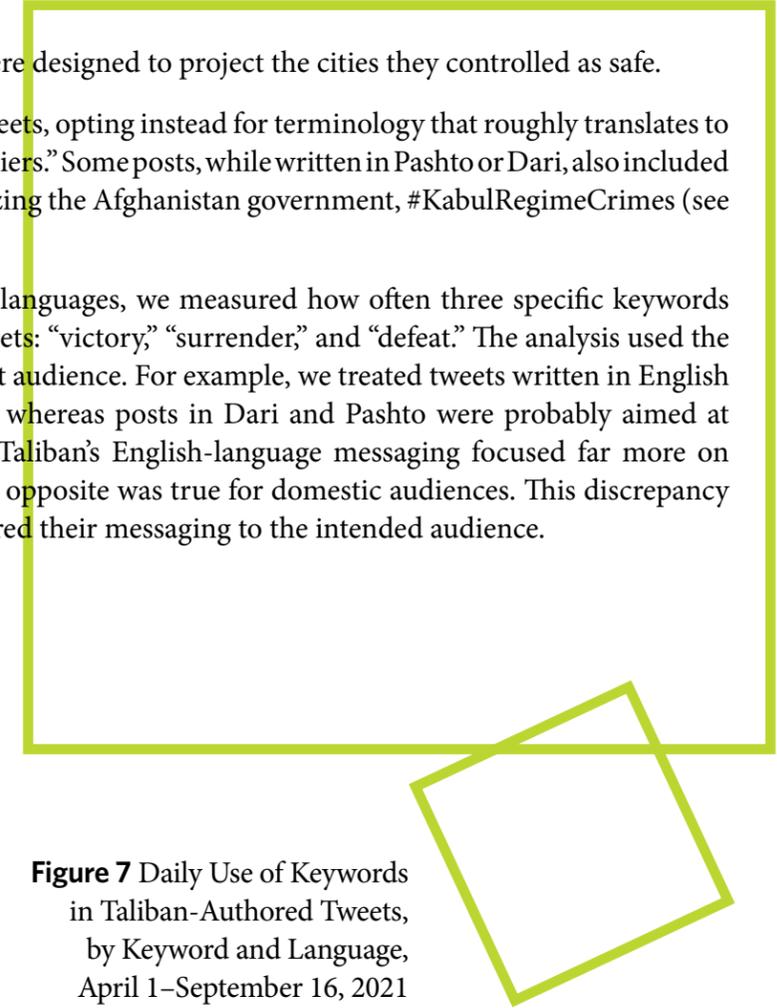


Figure 7 Daily Use of Keywords in Taliban-Authored Tweets, by Keyword and Language, April 1–September 16, 2021



¹⁸ See Bakhtar News Agency (2021).

III. TWITTER'S MODERATION

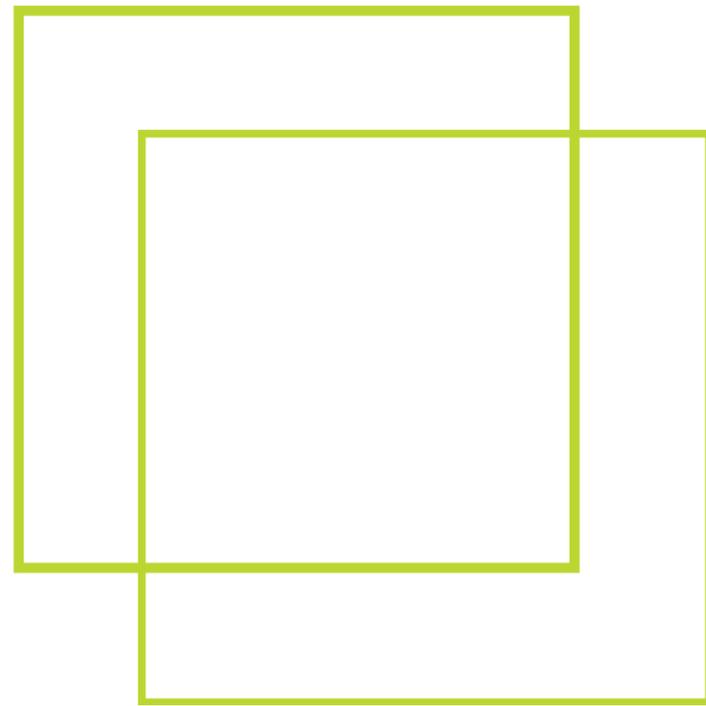


Figure 8 Daily Use of Keywords in Taliban-Authoring Tweets, by Keyword and Language, April 1–September 16, 2021



Note: The graphs show the daily number of retweets, replies, likes, and quote tweets elicited by Taliban-authored tweets.

Figure 8 illustrates how often audiences engaged with Taliban-authored content featuring the same keywords. Engagement comprises retweets, replies, likes, and quote tweets. Variations across languages may reflect the authors' different approaches to information campaigns. As may be expected, English-language content received comparatively less engagement than the other languages.

The Taliban's social media outreach campaign and supporting network of accounts appear to have relied heavily on Twitter from April to September 2021. That reliance may be linked to Twitter's moderation practices, as indicated by two key conclusions of our analysis:

- Twitter's moderation failed in the lead-up to the takeover of Afghanistan.
- Twitter appears to have profited from the Taliban's presence on its platform.

These findings are supported by evidence that the Taliban conducted nearly all their social media activities on Twitter. Had they been highly active on other platforms or attempted to migrate their audiences away from Twitter, the analysis would have shown that Taliban accounts frequently posted external links to redirect their audiences. As noted above, however, links to other social media platforms accounted for less than 1% of all Taliban-authored content.

The moderation efforts of other platforms may have reinforced the Taliban's reliance on Twitter. Our analysis indicates that links included in tweets to Facebook are no longer active. This deactivation may be the result of moderation efforts by Facebook, although it is possible that the Taliban made an organizational decision to remove the links.¹⁹

Twitter's Failed Moderation of Taliban Accounts

Twitter's policy on violent organizations, updated in October 2020, states: "under this policy, [users] can't affiliate with and promote the illicit activities of a terrorist organization or violent extremist group."²⁰ In the months leading up to the takeover of Kabul, however, Taliban-branded Twitter accounts provided hourly updates promoting the group's political activities and military victories.

The updates involved graphic videos and images of combat and wounded or dead individuals (see Image 26 in the Appendix). These posts violated Twitter's rules by showing graphic content, including photographs and videos of mutilated corpses. In particular, Twitter's "sensitive media policy" prohibited images depicting "gratuitous gore" and "graphic violence."²¹ Many posts also included calls for violence against the Afghan government and U.S. forces.

Our assessment of the broader Taliban support system of more than 126,000 accounts shows evidence that Twitter moderated only 49 accounts. In some of these cases, Twitter locked offending accounts and displayed the following suspension message: "Account is temporarily unavailable because it violates the Twitter media policy." As discussed in Section I and shown in Figure 2, the vast majority (83%) of Taliban-associated accounts were created before 2021, long before Twitter could claim that their content was permissible because the group had become the government of Afghanistan.²²

¹⁹ Facebook has stated that it would take action against Taliban-linked accounts and networks (BBC 2021).

²⁰ See Twitter (n.d.b).

²¹ See Twitter (n.d.a).

²² See Madhok (2021) and Nix and Bloomberg (2021).



While social media platforms often claim that networks of violent and dangerous actors can be challenging to moderate effectively, the core Taliban network was highly centralized, which would arguably have made Taliban content easier to monitor and regulate. As noted in Section I, more than two-thirds of all Taliban-authored tweets in the period under review came from only 20 accounts, while almost half (47%) of all tweets can be traced to only ten accounts. These findings suggest that Twitter devoted minimal effort to moderating this network, although a crackdown could potentially have had a significant impact.

Twitter's Monetization of the Taliban

Our research shows that Twitter placed ads paid for by U.S. and Canadian companies as sponsored tweets on the accounts of the Taliban news services, spokespersons, and senior leaders. Our investigation also exposes an apparent disconnect between Twitter's advertising arm and its content moderation unit. It shows that the company violated its stated policies by placing ads on accounts that had (or should have) been flagged as violating company policy because of posts containing violent and graphic content. A tweet posted on August 7, for example, contained an image of a decomposing corpse. As of 21 September, the tweet was still visible—and the account featured sponsored tweets. The account had been flagged by Twitter's automated moderation tool, as evidenced by a warning screen viewed by users upon visiting the account; nevertheless, it continued to feature ads (see Image 26 in the Appendix).

North American Twitter users saw ads from Amazon, Disney, McDonald's, and Royal Bank of Canada, among other companies.²³ Twitter states that promoted ads only “display on select profiles that fit the targeting credentials configured for a campaign,” which suggests that the company deliberately allowed advertisers to monetize the Taliban's audience.²⁴ More than 30% (19 of 63 accounts) of the core network of Taliban-branded accounts featured sponsored ads.

23 See Images 2–5 in the Appendix.

As of May 2022, these accounts remained active, attracting a growing number of followers and continuing to post content in violation of Twitter's stated policies. At this writing, the core Taliban network had a following of more than 3.3 million and Twitter continued to monetize the group's presence on the platform. Thirteen accounts were no longer available on Twitter, including four accounts that Twitter had suspended (as indicated by landing pages when reviewing the account URLs) and nine that no longer existed. It is unclear whether Twitter removed the accounts labeled “no longer exists”; Taliban users could have deactivated or renamed them. Two accounts were labeled “temporarily restricted,” such that users must click through to view the profile. One of these temporarily restricted accounts featured sponsored ads (see Image 7 in the Appendix), suggesting a lack of coordination between Twitter's moderation and advertising functions. Advertisements remained on the account profiles of 40% (25 out of 63) of the studied accounts, despite removals. A case in point is ongoing advertising (see Image 27) on the account of Anas Haqqani, leader of the Haqqani network, which the U.S. government designated a foreign terrorist organization in 2012.²⁵

We were unable to document how long sponsored ads appeared on Taliban accounts due to the limitations of Twitter's API. Instead, we manually viewed these accounts periodically to determine whether sponsored ads appeared on the 63 accounts.

In light of recent investigations indicating that Facebook made millions of dollars in ad revenue from disinformation and inadvertently funded clickbait actors through ad funding, platforms must consider this problem more seriously.²⁶ With support from the platforms, researchers and civil society can help identify ads on malicious actors' accounts. This could be an extension of existing programs focused on issues of ad placement, including Twitter's Ads Transparency Center and Facebook's Ad Library.

25 See USDoS (n.d.).

IMPOLICY IMPLICATIONS

The evidence presented in this report suggests that Twitter played a crucial role in the Taliban's takeover of Afghanistan. Their prolific and sophisticated use of Twitter was key to their influence operation, which allowed the group to project itself as the inevitable victor, while amplifying the failures of the Afghan government and its Western backers. The Taliban's social media prowess demonstrates how non-state armed groups can advance online influence campaigns targeting domestic and international audiences, particularly in the absence of effective content moderation.

Our findings strongly suggest that social media companies could do more to limit malicious actors' access to their platforms. Despite growing evidence that moderation strategies are insufficient, many platforms remain ill-equipped to deal with armed groups that, like the Taliban, have dedicated social media and propaganda teams. Moderation shortcomings may also reflect company priorities. Indeed, one recent investigation revealed that although more than 90% of Facebook's monthly active users live outside of the United States and Canada, 87% of the company's moderation efforts are focused on U.S.-oriented posts.²⁷ Another recent inquiry indicates that Facebook and Google are using ad revenue to fund destabilizing misinformation and influence operations around the world, further highlighting the urgent need to address the effectiveness of moderation tools and policies.²⁸

A better approach would involve tools and methods tailored to each conflict context, combined with a detailed understanding of the prominent armed groups and their propaganda strategies. One-time, blanket bans of groups such as the Taliban are unlikely to have the desired impact, as they can typically be circumvented.²⁹ Effective monitoring strategies require governments and social media companies to adapt faster than actors who violate applicable policies. Social media platforms are being weaponized, leading to a new arms race, one that is being lost by governments and social media companies. To ensure their moderation approaches are fit for purpose, we recommend that social media companies consider the following guidance:

- **Apply existing content moderation policies.** Twitter shied away from an outright ban of the Taliban in the aftermath of the group's takeover of the country. Instead, the company stated that it would remove Taliban posts that violated platform policies. We found repeated evidence that Twitter's moderation efforts failed to enforce company policies, permitting graphic images and videos—along with explicit calls to violence—to remain on the platform (see Images 26 in the Appendix). Twitter could have also applied its counter-radicalization policies, including the redirect method (which involves redirecting users to alternative content when they search for terms associated with extremism).³⁰ The company could also have deployed a range of approaches beyond removals, such as labeling accounts (as it does with Russian and Chinese state media

accounts) or placing labels on incendiary content (as it does with COVID and election disinformation).³¹

- **Increase coordination between internal teams.** Improving communication between advertising teams and units responsible for moderating harmful content can help enhance vetting processes and policy compliance. Better coordination could have prevented sponsored ads from appearing on Taliban accounts that were flagged for sharing graphic and violent imagery. Similarly, better coordination between moderation and advertising teams could have removed ads accounts that were temporarily restricted.

- **Move away from blocking offending content to targeting the actors and accounts producing it.** This shift in moderation techniques involves reducing the reliance on reactive strategies that respond to single inappropriate activities and instead increasing the capacity to track offenders and their networks. Developing and maintaining an actor-centered approach requires continuous input from relevant experts. By undertaking this transition, social media platforms can enhance their ability to prevent and stem the spread of harmful content before it reaches large audiences.

- **Stay one step ahead of targeted actors through ongoing monitoring and adaptation.** Current moderation efforts tend to be based on community standards and policies, often with the aim of managing violations rather than preventing or minimizing them. Given that offending groups typically find ways to circumvent rules and restrictions, social media companies can retain their edge by focusing on building—and sustaining—a continuous learning system with the capacity to monitor and adapt more quickly than these actors. At a minimum, this type of system requires interdisciplinary teams established in partnership with civil society organizations, governments, industry, and academia. Any automation efforts, guided by a database of keywords, would require regular updates from authorities in relevant fields, as well as experts on armed groups' social media strategies

- **Provide researchers with greater access to data.** Our ability to conduct this research reflects the accessibility of Twitter data through its publicly available API. While we found evidence of Taliban activity on other social media platforms (including Facebook, Instagram, and YouTube), we were limited in our ability to examine its scope. Social media companies that are serious about understanding the misuse and weaponization of their platforms by violent and extremist actors, must equip researchers with the tools required to quantify the breadth and extent of the problem. With support from the platforms, researchers and civil society can help identify ads on malicious actors' accounts. This work could help support programs that are already focused on issues related to ad placement, including Twitter's Ads Transparency Center and Facebook's Ad Library.

Since their takeover of Afghanistan in August 2021, the Taliban have consolidated their control. Under their rule, girls are banned from accessing education and women are required to wear burqas in public. The group has also kidnapped and tortured countless Afghans for being critical of their regime. The Taliban's exploitation of social media is a cautionary tale. Western governments had a clear security interest in Afghanistan throughout the two-decade-long conflict. If the Taliban did not warrant close moderation, what might this mean for the online landscape of other conflicts? How many other governments and armed groups may be exploiting social media in ways we have yet to discover?

27 See Tworek (2021).

28 See Hao (2021).

29 See Chakravorti (2021), Ghaffary (2021), and Khilji (2021).

30 See Greer and Ramalingam (2020).

31 See Twitter (n.d.d; n.d.e; n.d.f).

References

- Atiq, Sarah. 2021. "The Taliban Embrace Social Media: 'We Too Want to Change Perceptions.'" BBC News, September 6, 2021. <https://www.bbc.com/news/world-asia-58466939>.
- Bahar, Hazrat M. "Social Media and Disinformation in War Propaganda: How Afghan Government and the Taliban Use Twitter." *Media Asia* 47, nos. 1–2 (2020): 34–46.
- Bakhtar News Agency. 2021. "Air Force Helicopter Crashed in Helmand." July 31, 2021. <https://bakhtarnews.af/air-force-helicopter-crashed-in-helmand/>.
- Bauder, David, Michael Liedtke, and the Associated Press. 2021. "Whistleblower says Facebook routinely chose 'profit over safety' when it came to misinformation." *Fortune*, October 4, 2021. <https://fortune.com/2021/10/04/facebook-whistleblower-social-media-misinformation-hate-algorithm/>
- BBC. 2021. "Afghanistan: Facebook Continues Ban of Taliban-Related Content." BBC, August 17, 2021. <https://www.bbc.com/news/business-58239786>.
- Bernatis, Vincent. "The Taliban and Twitter: Tactical Reporting and Strategic Messaging." *Perspectives on Terrorism* 8, no. 6 (2014): 25–35.
- Brooking, Emerson T. 2021. "Before the Taliban Took Afghanistan, It Took the Internet." Atlantic Council Blog, August 26, 2021. <https://www.atlanticcouncil.org/blogs/new-atlanticist/before-the-taliban-took-afghanistan-it-took-the-internet/>.
- Butler, Gavin. 2021. "How Social Media Helped 'Taliban 2.0' Take Control of Afghanistan." VICE News, August 19, 2021. <https://www.vice.com/en/article/z3x5we/social-media-helped-taliban-control-afghanistan>.
- Chakravorti, Bhaskar. 2021. "Facebook's Taliban Ban Will Prove Costly for Afghans." *Foreign Policy*, August 26, 2021. <https://foreignpolicy.com/2021/08/26/taliban-facebook-ban-social-media-afghanistan/>.
- Culliford, Elizabeth. 2021. "YouTube Says It Does Not Allow Taliban-Affiliated Accounts." Reuters, August 17, 2021. <https://www.reuters.com/technology/youtube-says-it-bans-accounts-believed-be-owned-by-taliban-2021-08-17/>.
- Drissel, David. "Reframing the Taliban Insurgency in Afghanistan: New Communication and Mobilization Strategies for the Twitter Generation." *Behavioral Sciences of Terrorism and Political Aggression* 7, no. 2 (2015): 97–128.
- Frenkel, Sheera, and Ben Decker. 2021. "Taliban Ramp Up on Social Media, Defying Bans by the Platforms." *The New York Times*, August 18, 2021. <https://www.nytimes.com/2021/08/18/technology/taliban-social-media-bans.html>.
- Ghaffary, Shirin. 2021. "How Facebook, Twitter, and YouTube Are Handling the Taliban." *Vox*, August 18, 2021. <https://www.vox.com/recode/22630869/facebook-youtube-twitter-taliban-ban-social-media-afghanistan>.
- Greer, Ryan and Vidhya Ramalingam. 2020. "The Search for Extremism: Deploying the Redirect Method." Washington Institute for Near East Policy. February 27, 2020. <https://www.washingtoninstitute.org/policy-analysis/search-extremism-deploying-redirect-method>.
- Hao, Karen. 2021. "How Facebook and Google Fund Global Misinformation." *MIT Technology Review*, November 20, 2021. <https://www.technologyreview.com/2021/11/20/1039076/facebook-google-disinformation-clickbait/>.
- Hendrix, Justin. 2021. "Social Media and the Future of the Taliban." *Tech Policy Press*, August 20, 2021. <https://techpolicy.press/social-media-and-the-future-of-the-taliban/>.
- Hussaini, Salim, and Travis Morris. "The Taliban's Information War: The Tactical Use of Frames." *Journal of Information Warfare* 19, no. 4 (2020): 89–109.
- Johnson, Thomas H., Matthew DuPee, and Wali Shaaker. "The Taliban's Use of the Internet, Social Media Video, Radio Stations, and Graffiti." In *Taliban Narratives: The Use and Power of Stories in the Afghanistan Conflict*, edited by Thomas H. Johnson, 85–106. Oxford University Press, 2018. <http://dx.doi.org/10.1093/oso/9780190840600.003.0001>.
- Khilji, Usama. 2021. "Should Facebook Let the Taliban Post?" *Slate*, September 14, 2021. <https://slate.com/technology/2021/09/facebook-twitter-taliban-social-media-content-moderation.html>.
- Madhok, Diksha. 2021. "How Social Media Is Dealing with the Taliban Takeover." *CNN Business*, August 17, 2021. <https://www.cnn.com/2021/08/17/tech/facebook-twitter-taliban-hnk-intl/index.html>.
- Mehran, Weeda, Umniah Al Bayati, Matthew Mottet, and Anthony F. Lemieux. "Deep Analysis of Taliban Videos: Differential Use of Multimodal, Visual and Sonic Forms across Strategic Themes." *Studies in Conflict & Terrorism* (January 2021): 1–21. DOI: 10.1080/1057610X.2020.1866739.
- Mozur, Paul, and Zia ur-Rehman. 2021. "How the Taliban Turned Social Media Into a Tool for Control." *The New York Times*, August 20, 2021. <https://www.nytimes.com/2021/08/20/technology/afghanistan-taliban-social-media.html>.
- Nix, Naomi, and Bloomberg. 2021. "Facebook and Twitter Weigh Giving a Social-Media Megaphone to the Taliban." *Fortune*, August 21, 2021. <https://fortune.com/2021/08/21/facebook-twitter-social-media-taliban/>.
- Singer, Peter Warren and Emerson T. Brooking. *LikeWar: The Weaponization of Social Media*. Eamon Dolan Books, 2018.
- Timberg, Craig. 2021. "Taliban Websites Operating in Five Languages Go Dark." *The Washington Post*, August 20, 2021. <https://www.washingtonpost.com/technology/2021/08/20/taliban-internet-websites-twitter-facebook/>.
- Twitter. n.d.a. "Sensitive Media Policy." <https://help.twitter.com/en/rules-and-policies/media-policy>.
- . n.d.b. "Violent Organizations Policy." <https://help.twitter.com/en/rules-and-policies/violent-groups>.
- . n.d.c. "What Are Promoted Ads?" <https://business.twitter.com/en/help/overview/what-are-promoted-ads.html>.
- . n.d.d. "New Labels for Government and State-Affiliated Media Accounts." https://blog.twitter.com/en_us/topics/product/2020/new-labels-for-government-and-state-affiliated-media-accounts.
- . n.d.e. "COVID-19 Misleading Information Policy." <https://help.twitter.com/en/rules-and-policies/medical-misinformation-policy>.
- . n.d.f. "Civic Integrity Policy." <https://help.twitter.com/en/rules-and-policies/election-integrity-policy>.
- Tworek, Heidi. 2021. "Facebook's America-Centrism Is Now Plain for All to See." *Centre for International Governance Innovation*, October 4, 2021. <https://www.cigionline.org/articles/facebooks-america-centrism-is-now-plain-for-all-to-see/>.
- USDoS (United States Department of State). n.d. "Foreign Terrorist Organizations." Bureau of Counterterrorism. <https://www.state.gov/foreign-terrorist-organizations/>.
- Wagner, Meg, Melissa Macaya, Fernando Alfonso III, Melissa Mahtani, and Mike Hayes. 2021. "Biden Announces Plan to End America's Longest War." *CNN*, April 14, 2021. <https://www.cnn.com/politics/live-news/biden-afghanistan-announcement/index.html>.

APPENDIX

CORROBORATING EVIDENCE FROM TALIBAN ACCOUNTS



Image 1: Taliban’s declaration of victory after the capture of Kabul. A spokesperson of the Islamic Emirate of Afghanistan, the Taliban’s self-proclaimed state, declared on Twitter: “With the help of God, and the support of the nation, we are now in control of all parts of the country. We would like to congratulate our nation on this big achievement.” August 15, 2021.

Image 2: The account of Emran Khalil, the author and administrator of the Taliban’s Pashto language website, Alemarah. This post was tweeted on June 18, 2021. This screenshot was taken on August 26, 2021 by a U.S.-based user.*



* This image was altered by the CAIDAC team to protect the identity of the individuals depicted in the image.

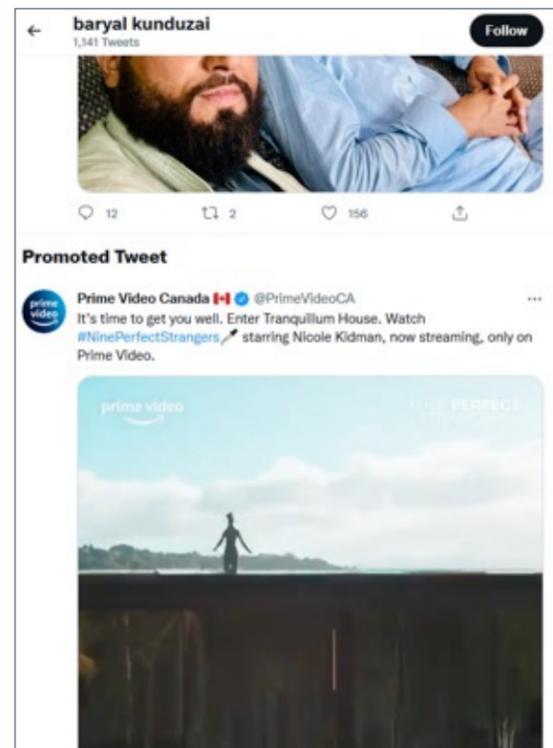
Image 3: The account of a Taliban member. This account would share links to other Taliban members' and supporters' accounts to encourage followers to view this content. This account also features a personal Facebook page in its account bio. Below is a sponsored ad from McDonald's, viewed on August 23, 2021 by a Canadian-based user. Sponsored ads are still featured on this account as of May 2022.



Image 6: Sponsored tweets for the Canadian Imperial Bank of Commerce featured on a Taliban account. September 7, 2021, captured by a Canadian user.



Image 7: Example of an account labeled "Caution: This profile may include potentially sensitive content" that nonetheless included paid advertising once users click through to the account. The account is explicitly Taliban-linked with Sponsored Tweets and graphic imagery in violation of Twitter's stated policy. The screenshot including the warning label was taken September 10, 2021 by a Canadian user. The second screenshot was taken on September 16, 2021 by a U.S. user but no warning was on the profile but an advertisement was present.



Images 4 and 5: Sponsored ads on two Taliban accounts from Amazon Prime Video Canada and Amazon Web Services. These ads were captured on September 7, 2021 by a Canadian user.





Image 8: Several accounts were found to be using an English-language hashtag (#taliban) in their bios to signal affiliation. They would also include links to the Taliban's official website (Alemarah) in different languages (in this case, Dari). Screenshot taken on August 23, 2021.



Image 9: Many accounts are overtly Taliban-branded, with Taliban flags in their profile or banner pictures. Screenshot taken on August 23, 2021.

Image 10: Tweets from Suhail Shaheen, an official Taliban spokesperson, on August 11, 2021. Screenshot taken on September 11, 2021.



Image 11: Tweets from Suhail Shaheen, an official Taliban spokesperson. The Tweets were published on June 20, 2021. Screenshot was taken on September 11, 2021 by a user in the US.

Image 13: The Taliban claimed to have shot down a helicopter. Afghan forces refuted this claim and stated they made an emergency landing due to technical problems.¹ Tweet posted on July 29, 2021. Screenshot taken September 11, 2021.

¹ <https://bakhtarnews.af/air-force-helicopter-crashed-in-helmand/>



9:59 AM · Jul 29, 2021 · Twitter Web App

Image 14: The Taliban posted a video of a father who claimed the Taliban saved his child and family. He claimed they live in Marjah District. Marjah District did not come under Taliban's control until July 21, 2021. Translation correction: The first sentence, "The whole nation owes the Taliban". July 13th, 2021. Screenshot taken on September 21, 2021

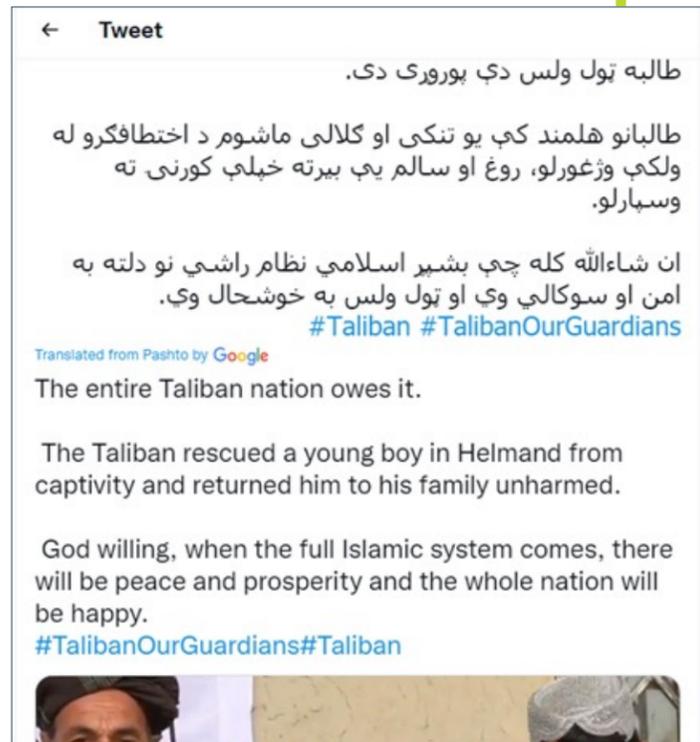


Image 15: Tweet from Bilal Karimi, an official member of the Taliban, criticizing the Afghan government. Tweeted on July 14, 2021. Screenshot on September 10, 2021.



Image 16: Another tweet from Bilal Karimi, an official member of the Taliban, criticizing the Afghan government. Tweeted on July 31, 2021. Screenshot on September 10, 2021.



Image 17: The account of the administrator responsible for the English-language version of the Taliban's official website, Alemarah. His profile featured Sponsored Tweets. Tweeted on September 16, 2021. Screenshot taken on September 17, 2021. He frequently mentioned international humanitarian and human rights organizations in Tweets criticizing U.S. forces and the Afghan government.



Image 18: The account of the administrator overseeing the English-language version of the Taliban's official website, Alemarah. His profile featured Sponsored Tweets. Tweeted on August 24, 2021. Screenshot taken on August 24, 2021. He frequently mentioned international humanitarian and human rights organizations in Tweets criticizing U.S. forces and the Afghan government.



Image 19: The Taliban highlighted actions that would qualify as war crimes under International Humanitarian Law attributed to the Afghan government. July 26, 2021. Screenshot captured on September 10, 2021.



Image 20: The Taliban accused the Afghan government of killing a large number of civilians. On the same date, the Afghan government reported heavy Taliban casualties.² August 1st, 2021. Screenshot taken on September 10, 2021.



² <https://bakhtarnews.af/heavy-casualties-inflicted-to-taliban-in-helmand-recent-clashes/>; <https://bakhtarnews.af/forty-nine-taliban-killed-in-helmand-operation/>

Image 21: The Taliban reported the defection of Afghan soldiers. July 26, 2021. Screenshot taken on September 11, 2021.*



* This image was altered by the CAIDAC team to protect the identity of the individuals depicted in the image.

Image 23: The Taliban announced the defection of Afghan soldiers. August 2, 2021. Screenshot taken September 11, 2021.



Image 22: The Taliban announce the defection of Afghan soldiers. August 2, 2021. Screenshot taken on September 11, 2021.*

* This image was altered by the CAIDAC team to protect the identity of the individuals depicted in the image.

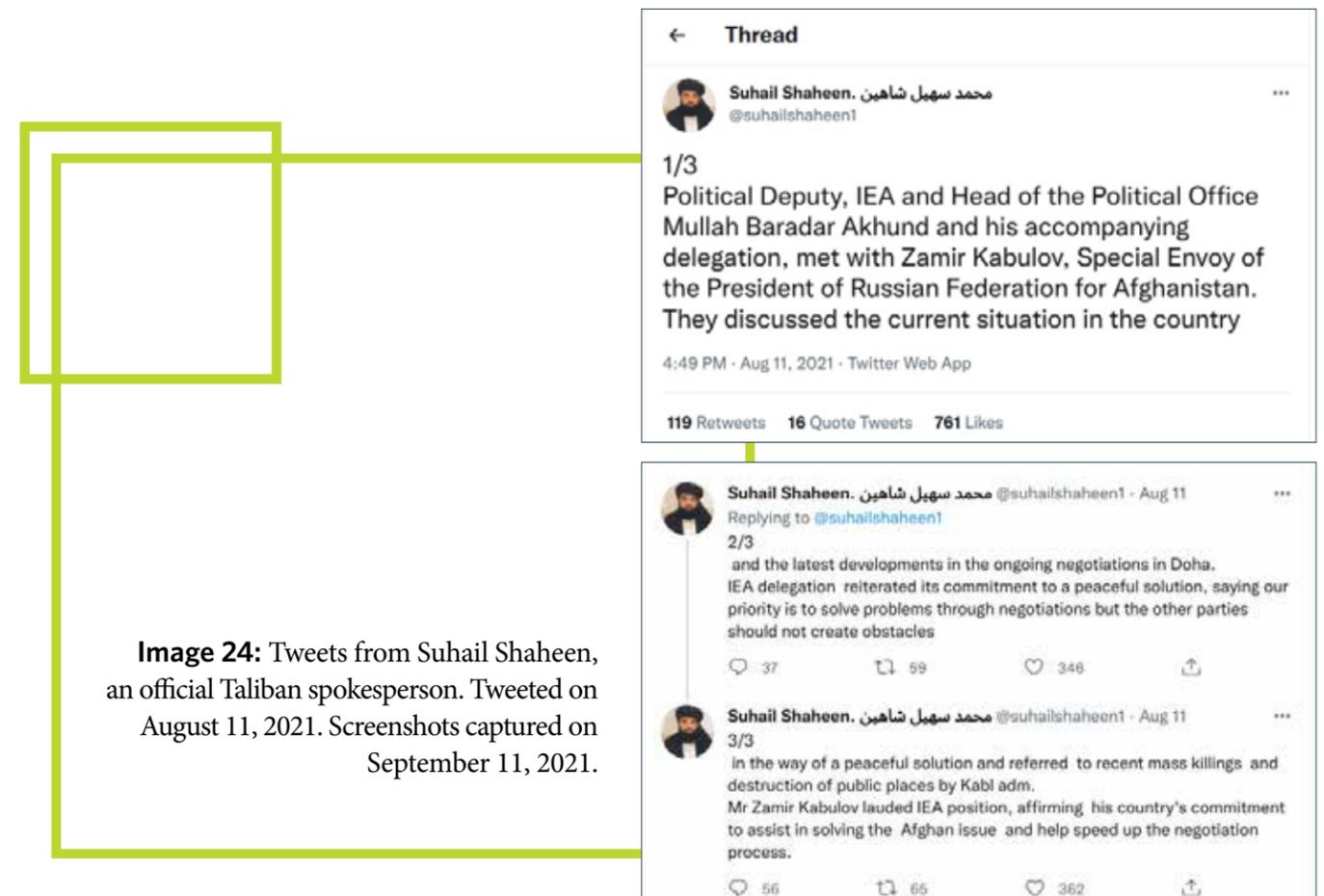


Image 24: Tweets from Suhail Shaheen, an official Taliban spokesperson. Tweeted on August 11, 2021. Screenshots captured on September 11, 2021.

Image 25: Tweets from Suhail Shaheen, an official Taliban spokesperson. Tweeted on July 31, 2021. Screenshots collected on September 11, 2021.



Image 26: A Taliban-claimed account with Sponsored Tweets and graphic imagery in violation of Twitter's stated policy. Screenshot on August 23, 2021.*

Image 27: The account of Anas Haqqani, leader of the Haqqani Network, featuring sponsored ads. The screenshot was taken on June 5, 2022.



* This image was altered by the CAIDAC team to protect the identity of the individuals depicted in the image.

Centre for Artificial Intelligence, Data, and Conflict

CAIDAC's goal is to create a global, state-of-the-art platform to share human-in-the-loop AI tools, labels, models, and algorithms to capture social media's transformation of conflict, political violence, and war.

Its mission was born from our experience living in communities affected by violence. We watched as social media transformed conflict and also provided unprecedented documentation of events on the ground. And yet, the methods and tools available to researchers and humanitarians to study conflict have not undergone a similar revolution. CAIDAC's researchers are motivated by the real-world impact of our research on practice and policy. We strive to be a place of innovation: a hub for individuals with different skill sets and backgrounds to address consequential problems they could not attempt alone. We value humility, unusual perspectives, and a desire to change the world (for the better). CAIDAC was co-founded by Laura Courchesne, Brian McQuinn, Cody Buntain, Denilson Barbosa, and Matthew Taylor.

Contact: admin@tracesofconflict.com

