



Fear, Lies and Lucre: How Criminal Groups Weaponise Social Media in Mexico

Crisis Group Latin America Briefing N°50

Mexico City/Bogotá/New York/Brussels, 31 January 2024

What's new? Mexico's criminal groups use social media to garner popular support, denigrate rivals, glorify narco-culture and coordinate violence. Social media is also crucial to providing timely information about flare-ups of violence, particularly since journalists face major threats to their safety, which heavily circumscribes their ability to report from many crime-affected municipalities.

Why does it matter? These criminal groups are recruiting and spreading disinformation online, making them stronger and creating a glut of unverified information that puts civilians at greater risk. Platforms have struggled to respond appropriately.

What should be done? Platforms should boost resources for online monitoring, especially when violence is spiking. Given social media's importance in disseminating information, platforms should modify their algorithms to demote posts supporting criminal groups and work with civil society to identify trusted accounts, including anonymous ones. Mexico's government should also invest in protecting journalists.

I. Overview

In Mexico, social media outlets have struggled to manage criminal groups' use of their platforms. Mexican syndicates post content that often glorifies breaking the law, promotes violence and extends their own reach while undermining rivals. Social media companies face a number of challenges in responsibly policing this content, particularly reporting from anonymous citizen journalists, which is often hard to verify. To prevent inaccurate information and propaganda from fuelling conflict and endangering civilians, platforms need to strengthen their capacity to moderate content and promote reliable sources. This task is all the more important because social media is fast becoming the dominant source of grassroots information on the latest episodes of violence from areas that are too dangerous for professional journalists to work in. Social media providers should work more closely with local civil society to develop new tools for keeping criminal groups offline, while the Mexican government should redouble efforts to protect journalists, thus helping ensure that the public has reliable sources of information.

Criminal groups have used popular platforms like Facebook, Twitter, TikTok, YouTube and Telegram, as well as Encrypted Messaging Apps (EMAs) such as WhatsApp, Inbox and Messenger, to turn their local feuds into digital fights. They exploit accounts

to threaten their foes, attempt to win over citizens and lure new recruits through glamourised depictions of life beyond the law. When the Jalisco New Generation Cartel (CJNG) and the United Cartels – an umbrella group encompassing various local outfits – battled for Michoacán state, their online threats against one another peaked with the apex of their fighting in 2021.

Platforms have notched some successes in reducing the amount of graphic violence on their sites, but thorny issues remain. Clamping down on criminal groups exploiting social media while not infringing on freedom of expression and information is tough. There is often a fine line between information and propaganda, or between online activity by criminal groups and more innocuous posting by ordinary citizens making pop-culture drug references. The rapid fragmentation and rebranding of criminal groups makes policies aimed at removing certain outfits and moderating their content – as most platforms attempt to do – an uphill battle.

Because journalists can no longer safely report from swathes of Mexico, social media has become a more influential news source, which in turn raises a series of challenges for operators of social media sites. A range of voices – from citizen journalists to sometimes sensationalist “narcoblogs” to the illegal outfits themselves – have sought to fill and sometimes exploit for their own ends the information vacuum in crime-affected parts of Mexico. Within this information ecosystem, social media accounts run by anonymous individuals documenting crime attempt to provide information to the public. Because they operate anonymously, however, it can be difficult for social media companies to discern which accounts are operated by citizen journalists and which are orchestrated by cartel members.

Online and offline efforts will be essential to stem the rising criminal use and misuse of social media. Platforms should strengthen their content moderation and work with local and regional civil society to verify trusted accounts – including those of anonymous citizen journalists – with the aim of “boosting” their content higher in users’ feeds. At the same time, social media firms should modify their algorithms to downgrade posts supporting criminal groups and expand shadow banning of criminal groups, meaning that users searching for these groups by name will not uncover any results. Lastly, the government of President Andrés Manuel López Obrador should pursue stronger measures to protect journalists so that crime-affected areas do not remain news deserts and more reliable information is available to people living there and throughout Mexico.

II. Criminal Groups Online

A. *A Brief History of Crime and Social Media in Mexico*

Criminal groups in Mexico first made systematic use of social media in the late 2000s, when they started to post videos of executions and interrogations on YouTube.¹ One shocking video in 2007 appeared to show the beating and beheading of a member of the crime group Los Zetas.² Uploaded videos have also depicted the torture of civil-

¹ Mica Rosenberg, “Narcotráfico México usa violentos videos para sembrar terror”, Reuters, 4 August 2010; and Rick Jervis, “YouTube riddled with drug cartel videos, messages”, ABC News, 9 April 2009.

² “Web video shows Mexico man’s beheading”, NBC News, 1 April 2007.

ians. In 2010 and 2011, a series of posts showed members of the Sinaloa Cartel beating the brother of a prosecutor and forcing him at gunpoint to “confess” his sister’s supposed alliance with the Juárez Cartel.³

In the mid-2010s, after the Islamic State of Iraq and Syria (ISIS) made vigorous use of social media to spread jihadist propaganda, social media platforms initiated a crackdown on graphic imagery and violent groups.⁴ ISIS responded by moving incendiary posts from its own “branded” accounts and channels on social media to “unofficial” accounts run by its members and supporters. As a result, criminal groups were able to continue disseminating their content on social media platforms. Mexico’s criminal groups mimicked that strategy, and communication from illegal outfits became more decentralised.⁵

In parallel, Mexican criminal groups have splintered and multiplied over the past decade or so. As Crisis Group has reported, the number of criminal groups in Mexico more than doubled in the last decade, going from 76 in 2010 – the first full year of data – to 205 in late 2020.⁶ This fragmentation is reflected online, with criminal groups and their members relying heavily on the work of members and supporters to provide platforms and boost profiles on social media. One anonymous citizen journalist estimated that half of all *sicarios* (paid hit men) have an active social media profile.⁷

Criminal groups in Mexico are most commonly active on Facebook, which has more than 110 million users in the country of 128 million, and also offers Messenger, an encrypted messaging service, which has more than 96 million users in the country.⁸ Meta is the parent company of both platforms. An internal report by Facebook researchers from 2021, pieces of which were read to Crisis Group, stated that “criminal organisations blatantly misuse our platform for coordination of offline harm, recruitment, promotion of narcoculture and trading drugs”.⁹ Pages found by Crisis Group on the platform clearly display their criminal affiliation: one profile said, under education, “Studied killing [homophobic slur] at SINALOA CARTEL”. Many of these pages include pictures of men holding guns and in combat gear, their faces obscured by masks or emojis. Other accounts do not display a clear criminal affiliation but frequently post information that favours a particular group. Some dedicated pages supposedly sponsored by the crime groups themselves remain, though many have been removed.¹⁰

³ Tracy Wilkinson, “Kidnapped Chihuahua attorney found dead”, *Los Angeles Times*, 6 November 2010, and “Third video of Prosecutor’s brother depicts torture”, *Borderland Beat*, 12 February 2011.

⁴ Ryan Broderick and Ellie Hall, “Tech platforms obliterated ISIS online. They could use the same tools on white nationalism”, *Buzzfeed*, 20 March 2019; Rebecca Heilweil, “YouTube says it’s better at removing videos that violate its rules, but those rules are in flux”, *Vox*, 6 April 2021.

⁵ Crisis Group telephone interview, Facebook employee, April 2021. “Plenty of people make pages for the cartels, but I’m sure the leaders themselves consider this a silly thing”. See also Nilda García, *Mexico’s Drug War and Criminal Networks: The Dark Side of Social Media* (Abingdon, 2020).

⁶ Crisis Group Commentary, “Crime in Pieces: The Effects of Mexico’s ‘War on Drugs’, Explained”, April 2022.

⁷ Crisis Group telephone interview, narcoblogger, September 2021.

⁸ “Facebook Users in Mexico”, NapoleonCat, June 2023.

⁹ Crisis Group interview, Facebook employee, August 2021.

¹⁰ Facebook has a system in place, in particular the Dangerous Organizations and Individuals policy, for identifying and censoring violent criminal groups, but it consistently fails.

But social media's relevance to criminal activity in Mexico extends beyond its uses as a propaganda platform for criminal groups. To its credit, social media has also played a critical role in documenting Mexico's drug war. Citizens use digital platforms to provide information about on-the-ground incidents involving criminal groups, including through what are referred to as "narcoblogs". Narcobloggers, anonymous citizen journalism websites and social media pages aimed at documenting violence, often publish or aggregate extremely violent and graphic content, taken from anonymous sources or from accounts close to criminal groups.

As Crisis Group has reported, around 80 municipalities in states such as Guanajuato, Guerrero, Zacatecas and Michoacán suffer extremely high rates of murder and other serious crimes.¹¹ Mainstream journalism in these crime-affected parts of the country is scant because reporters are under lethal threat if they report on criminal groups.¹² Indeed, Mexico is the most dangerous country in Latin America for journalists, with fifteen killed in 2022 alone.¹³ After the murder of two staff members in 2010, one newspaper in Juárez published an open letter asking criminal groups for clearer content instructions: "We ask you to explain what you want from us, what we should publish or not publish, so we know what to expect".¹⁴

While the decline in mainstream media coverage of dangerous areas has shifted attention to narcoblogs, relying on these outlets to provide information about criminal movements generates risks of its own. Critics, among them other news sources, have accused these blogs of serving as mouthpieces and amplifiers of criminal propaganda, arguing that violent content only serves to glorify the lifestyles criminal group members pursue – not to mention the violence they perpetuate.¹⁵ Narcobloggers counter that graphic images of violence are newsworthy and should be shared since they depict the brutal reality of criminal violence.¹⁶ The nature of the relationship between criminal groups and narcobloggers varies: one administrator of a Twitter page dedicated to the drug war said they "highly doubted" anyone from a criminal group would provide them information; another said they frequently receive tips from criminal groups.¹⁷

Like members of the press, citizen journalists face threats to their safety. In 2014, María del Rosario Fuentes Rubio, a contributor to the narcoblog *Courage for Tamau-*

¹¹ Crisis Group Latin America Report N°99, *Mexico's Forgotten Mayors: The Role of Local Government in Fighting Crime*, 23 June 2023.

¹² Crisis Group Commentary, "Crime in Pieces: The Effects of Mexico's 'War on Drugs', Explained", op. cit. Oliver Darcy, "Anxiety and anger grip press corps in Mexico after spate of murders", CNN, 27 January 2022.

¹³ "2022 has been the deadliest year on record for Mexican journalists", PBS, 17 December 2022. The number killed in 2023 is variously estimated at anywhere from five to ten. Luis Carlos Sáinz, "10 periodistas asesinados en 2023 en México", *Semanario Zeta*, 9 January 2024.

¹⁴ "Qué quieren de nosotros? pregunta diario mexicano a narcotraficantes", *El Universo*, 19 September 2010.

¹⁵ "Some have called me a *puntero* [watchman] or a *halcón* [scout] because I show signs and posters that organised crime leaves behind". Crisis Group correspondence, narcoblog account, September 2021. See also Hector Amaya, *Trafficking: Narcoculture in Mexico and the United States* (Durham, NC, 2020); and "Blog del Narco gana fama y genera fuerte polémica", *La Tercera*, 14 August 2010.

¹⁶ One narcoblog Twitter account compared censorship of narco-messages in the Mexican press to Russian censorship. See post by @just_some_dOOd, 10:14pm, 14 March 2020.

¹⁷ Crisis Group telephone interviews, narcoblog accounts, September 2021.

lipas, was tortured and murdered. The group responsible got into her Twitter account, posting: “Today my life has reached its end” and “I can only tell you to not make the same mistake I did”. The year before, a criminal group had offered a reward for information about who was behind the anonymous blog.¹⁸ One narcoblog account administrator reported facing kidnapping threats.¹⁹ Another anonymous blogger alleged that members of the state security forces “were charged with finding your location and bothering you to get you to stop uploading information against them”.²⁰ The bloggers nevertheless say they continue their work because they feel it is crucial to provide accurate information about the human costs of the war on drugs and the spread of criminal violence it has given rise to.²¹ Social media, then, has served both as a propaganda tool for violent groups and as a method of information sharing about criminal activity and its causes.

B. *Platform Rules on Criminal Groups*

Major social media platforms have not been passive in the face of their exploitation by criminal organisations. Each maintains a set of rules under which they practice content moderation. Most public platforms seek to counteract criminal content in three ways: deactivating accounts related to groups deemed violent or dangerous; removing specific posts related to violence; and cancelling accounts that repeatedly violate content rules. EMAs like WhatsApp and Signal, in contrast, very rarely ban users.

Beyond moderating content, major social media companies have policies for keeping violent groups off their sites. Facebook, for example, keeps a tiered “dangerous organisations” list. Criminal groups fall under “Tier 1” alongside terrorist and hate groups: consistent with this status they are not allowed to have a presence on the platform and other users are not permitted to praise these groups.²² The platform formerly called Twitter, now known as X, banned violent organisations that have a “shared purpose” and have “systematically targeted civilians” according to the terms of its 2020 policy.²³ Since Elon Musk’s takeover of the platform in October 2022, these terms of use have broadly remained in place, although the reorganisation of the company led to the departure of many of those responsible for constructing and enforcing the policy toward violent groups.²⁴

¹⁸ “Valor por Tamaulipas contributor tortured and killed”, Justice in Mexico, 29 October 2014.

¹⁹ Crisis Group telephone interview, narcoblog account, September 2021.

²⁰ Crisis Group telephone interview, narcoblog account, September 2021.

²¹ Jane Esberg, “More than Cartels: Counting Mexico’s Crime Rings”, Crisis Group Commentary, 8 May 2020.

²² Sam Biddle, “Revealed: Facebook’s secret blacklist of ‘dangerous individuals and organizations’”, *The Intercept*, 12 October 2021.

²³ “Violent organizations policy”, Twitter, October 2020. This briefing will refer to Twitter when describing interviews or incidents occurring before Musk renamed the platform in July 2023.

²⁴ According to the terms published by X, “There is no place on X for violent and hateful entities, including (but not limited to) terrorist organizations, violent extremist groups, perpetrators of violent attacks, or individuals who affiliate with and promote their illicit activities. The violence and hate these entities engage in and/or promote jeopardizes the physical safety of those targeted”. See “Violent and hateful entities policy”, X Help Center. Elon Musk dissolved Twitter’s Trust & Safety Council and axed the human rights and public policy teams, leaving an opaque supervision system in their place. See “Musk’s Twitter has dissolved its Trust and Safety Council”, NPR, 12 December

Other social media organisations approach the challenge in varied ways. TikTok, for its part, does not allow criminal groups or content that supports these groups.²⁵ YouTube removes content produced by violent criminal organisations.²⁶ In some cases, groups are “shadow-banned”, meaning that their names cannot be searched for directly. By the end of 2021, Instagram, Facebook and TikTok had all instituted a shadow ban on terms linked to the Jalisco New Generation Cartel, one of Mexico’s largest criminal groups. Searching for profiles containing “CJNG” on Facebook, for example, returned dozens of results in the spring of 2021 but nothing in 2022.

Content may also be removed for violating a specific platform policy related to harassment or violence. These posts are generally brought to the attention of content moderators through reviews conducted using artificial intelligence or by users of the site. Social media platforms generally remove particularly graphic content, such as posts or videos that depict the moment of death, though X has had more relaxed rules on showing violence than other platforms.²⁷ Facebook and X provide the person posting the option of placing content behind a warning screen. TikTok may suspend or ban an account after any display of gratuitous violence, but the other platforms examined generally remove accounts only following repeated violations of content rules. Even so, platforms make an exception for violent posts that are educational or newsworthy, although each one has different parameters of what it will allow.²⁸

Platforms are more inclined to monitor and remove harmful content related to criminal groups and work with Mexican authorities during periods of potentially heightened tensions. That is especially true during election seasons. Meta, for example, increased monitoring of its various platforms – which include Facebook and Messenger, as noted, as well as Instagram and WhatsApp – in advance of Mexico’s 2021 elections due to concerns about the safety of candidates.²⁹ More generally, Twitter and Facebook both reported coordinating with law enforcement, including Mexico’s prosecutor’s office and the National Guard, in an attempt to both improve their content moderation and provide information about criminal activities to authorities.

2022; and Taylor Hatmaker, “Elon Musk just axed key Twitter teams like human rights, accessibility, AI ethics and curation”, 4 November 2022.

²⁵ Crisis Group telephone interview, TikTok employees, 28 October 2022. “Community Guidelines”, TikTok.

²⁶ “Violent criminal organization policy”, YouTube.

²⁷ “Twitter does proactive image detection, but posting nudity and gore is not a suspendable violation. Nudity is not banned and gore is not banned. We have policies around particular types of gratuitous gore, like moments of death and injury to protect victims. Really graphic dismemberments, beheadings ... we’re not proactively surfacing that, but unlike other platforms it’s not always categorically banned”. Crisis Group telephone interview, Twitter employee, October 2021. Users notice this difference: “Twitter and Telegram channels contain more information than other social media platforms. The new filters on Facebook are more aggressive and it is easier to be censored”. Crisis Group telephone interview, narcoblog account, September 2021. While this comment preceded Musk’s takeover of Twitter, social media researchers have found evidence that if anything rules on content have got looser. See Sheera Frenkel and Kate Conger, “Hate speech’s rise on Twitter is unprecedented, researchers find”, *The New York Times*, 2 December 2022.

²⁸ “Violent and graphic content”, Meta.

²⁹ Crisis Group interviews, Facebook employees, August 2021. On the criminal violence surrounding the 2021 polls, see Crisis Group Latin America Report N°89, *Electoral Violence and Illicit Influence in Mexico’s Hot Land*, 2 June 2021.

Facebook even launched a chat bot for Mexico City residents to report crime.³⁰ Platforms reported working with civil society partners to ensure they were updated about potential criminal activity.³¹

That said, long-term challenges remain so long as Mexican criminal groups maintain their presence on social media. They find means of skirting the rules in order to keep posting graphic imagery and other forms of propaganda. Meanwhile, private, encrypted messaging tools – including Messenger and WhatsApp – are widely used for coordination among members of criminal groups.

III. How Criminal Groups Use Social Media

Criminal groups use social media to intimidate rivals, win public approval and attract recruits. Torture and execution videos are still the mainstay of their social media strategy, but they also post about their charitable work and the ostensibly glamorous lifestyles of their operatives.

A. Threatening Rivals and Evading Moderators

Intimidating rival groups and civilians remains at the core of criminal enterprises' use of social media. To this end, they post violent imagery, including photos and videos of beheadings, beatings and torture. This type of content tends to proliferate during conflicts over territory, such as the escalation of fighting in Michoacán between the Jalisco New Generation Cartel and United Cartels in 2020-2021, which were at the time the two biggest criminal groups operating in that region.³² Threats are often very direct. One purported Jalisco member posted about a rival leader: "Don't worry, your death won't be so easy. You're going to suffer because you owe a lot of innocent people".³³

Social media is frequently used to "out" rival cartels' members, providing pictures and names of individuals allegedly working with adversaries.³⁴ Criminal groups also use platforms to demonstrate their military capacity and send threats. In at least one case, the Jalisco New Generation Cartel claimed to take over an Instagram account purportedly belonging to a member of a rival group called the Viagras following his death. Ominous posts soon cropped up on the feed, asserting supremacy over rivals in the fight for territorial control: "This account now belongs to the Jalisco New Generation Cartel. This is a message for everyone who takes the M2, Mencho and all

³⁰ Crisis Group telephone interviews, Twitter employees, October 2021; and Crisis Group telephone interviews, Facebook employees, August 2021. Crisis Group received no reply to inquiries with X about whether this practice has continued following the company's name change.

³¹ Crisis Group telephone interviews, Facebook employees, August 2021; Crisis Group telephone interviews, Twitter employees, October 2021.

³² On 31 October 2020, for example, a page allegedly run by members of the United Cartels posted an image of the mutilated bodies of purported CJNG members. For more on the fighting in Michoacán, see Falko Ernst, "On the Front Lines of the Hot Land: Mexico's Incessant Conflict", Crisis Group Commentary, 26 April 2022.

³³ Facebook post by purported CJNG member, 4 May 2021.

³⁴ One Facebook post showed several pictures of a woman, alongside the text: "This is the woman who is collecting the *piso* [extortion money] now. ... Her name is [name]". Facebook post from 11 April 2021.

the commanders as a joke. We will not take this lack of respect. This will happen to anyone who does not respect Señor Mencho. Good night”.³⁵

Criminal groups also threaten civilians, particularly those reporting on violence. In a video that circulated on social media, armed men read a message purportedly written by El Mencho and directed at journalist Azucena Uresti, who reports on the conflict in Michoacán. The statement declared: “I assure you that wherever you are I will find you and I will make you eat your words even if they accuse me of femicide”.³⁶ Sometimes, threats are intended to terrify a wider audience. The image of a body in a car riddled with bullet holes was accompanied by text reading: “Here and in all Guanajuato rules Señor Marro. They get in line or we get them in line”.³⁷

Though these types of posts clearly violate content rules, some of those reviewed by Crisis Group have remained online for months. One reason is flaws in content moderation. Most posts that are removed are first flagged by users or identified through artificial intelligence.³⁸ Users may, however, be unfamiliar with platform rules or how to flag content, and they may fear reporting on cartel posts online. Artificial intelligence-based detection of violence, meanwhile, is highly imperfect: it is difficult, for example, to differentiate between a screenshot from an action film and an image of real violence.³⁹ Moderators and automated control mechanisms may be under-resourced and not grasp the signs of criminal messaging. For example, groups often do not mention their names in posts so as to prevent easy recognition. Instead, they use emojis to identify themselves: a purple devil for the CJNG, pills for the Viagras and pickaxes for the Santa Rosa de Lima Cartel.⁴⁰

Moreover, criminal groups are often intentional and strategic about skirting content moderation. In 2021, the United Cartels claimed to be running multiple Facebook pages, now all removed, which cross-posted from one to another. The page “Carteles Unidos” largely just reposted content from “Defensores Michoacanos”, for example. Even if the latter page was suspended for violating Facebook’s community standards, the former might not face consequences. “They have paused four of our accounts for 40 days”, a page administrator wrote. “We will make more pages so you all know the truth”.⁴¹

³⁵ Instagram story, 8 April 2021, reported on a Mexican crime-focused Telegram channel on 12 April 2021. Instagram stories is an app feature allowing posters to share photos and videos that disappear within 24 hours. El Mencho (Nemesio Rubén Oseguera Cervantes) and M2 (Misael Torres Urrea) were two high-ranking criminals. At the time of the message, M2 was a lieutenant in the CJNG under El Mencho’s command; he eventually betrayed Oseguera, joined the group Los Viagras and was killed in 2022. “Ofrecen 5 millones de dólares por información para el arresto de hijastro del líder de cartel mexicano”, *El Universo*, 1 December 2021. “Qué se sabe sobre el asesinato del M2, el narcotraficante perteneciente al CJNG que traicionó a El Mencho”, *Prensa Libre*, February 11, 2022.

³⁶ “Amenaza Cártel Jalisco a la periodista Azucena Uresti: gobierno anuncia medidas”, *Aristegui Noticias*, 9 August 2021.

³⁷ Facebook post from 1 June 2020. Señor Marro is the leader of the Santa Rosa de Lima Cartel.

³⁸ Nafia Chowdhury, “Automated Content Moderation: A Primer”, Stanford Internet Observatory, 2022.

³⁹ Crisis Group interview, high-ranking Twitter employee, October 2021.

⁴⁰ Based on Crisis Group social media research. See, eg, Facebook post by purported members of the Santa Rosa de Lima Cartel, 1 June 2020; Facebook post by person alleging affiliation with the Viagras, 10 January 2021; TikTok video by alleged CJNG member, 16 April 2021.

⁴¹ Facebook comment by purported United Cartels page, 20 April 2021.

As one Facebook employee told Crisis Group, criminal groups post the most violent content on accounts they know will be short-lived: “Cartels use a ton of burner accounts that post the really bad stuff. They just have a couple of posts in their name over a short period of time and then not again”.⁴² Purported affiliates of the group can then repost this content, until the image is flagged and removed by moderators. Images may also be shared to WhatsApp, which is encrypted, or Telegram, which has been less aggressive about removing violent content than other platforms.⁴³

Videos and images can also find a second life through news and citizen journalism pages, which have greater freedom to post violent content so long as it is considered newsworthy. Narcobloggers frequently post graphic pictures alongside criminal threats – particularly *narcomantas* banners printed with threatening messages – left at the scenes of murders.⁴⁴ That said, despite offering greater leeway to posts deemed to have news interest, social media platforms will still take down posts that cross certain lines. Bloggers have strategies for keeping that from happening. Much like the strategies used by criminal groups and described above, they operate on multiple platforms to take advantage of variations in rules and enforcement. One blogger, who has most of his followers on Facebook, sometimes linked to his Twitter account because of its more relaxed rules on violence: “assassins dead and abandoned in the hills, the clearest images can be consulted on Twitter for those interested”.⁴⁵

Because they operate anonymously and sometimes post similar content, it can be difficult for social media companies to discern which accounts are operated by citizen journalists and which are orchestrated by cartel members.⁴⁶ Distinguishing propaganda from grassroots reporting becomes very difficult in the absence of trusted journalists, as several employees at Twitter told Crisis Group. “It’s difficult to evaluate whether an account is really run by a cartel or is just a local blog”, one observed.⁴⁷ Citizen journalists often rely on reporting from the ground that is not always verifiable. Even if a citizen journalist is providing truthful reporting, he or she may be accused of links to one side of a conflict or another.⁴⁸ Furthermore, researchers have noted that the content of some narcoblogs and posts on other platforms sometimes

⁴² Crisis Group interview, Facebook employee, April 2021.

⁴³ Telegram is an encrypted messaging app, which means that the content of the posts cannot be read by people who are in the chain of communication. It differs from similar apps in that it has “channels” that allow many people to join. Those channels have a wider audience, and yet the content is not visible to moderators. For details on the perils of encrypted messaging apps, see “Are private messaging apps the next misinformation hot spot?”, *The New York Times*, 30 June 2021. “Telegram gains users but struggles to remove violent content”, Trust and Safety Foundation, 2021.

⁴⁴ See, for example, “Ya empezó la narcoguerra contra el Cártel del Noreste, en narcomantas exhibieron a miembros (video)”, *El Blog del Narco*, 22 April 2022.

⁴⁵ Facebook post from narcoblogger allegedly aligned with the CJNG, 14 April 2021.

⁴⁶ Crisis Group interview, high-ranking Twitter employee, October 2021.

⁴⁷ Crisis Group telephone interviews, Twitter employees, October 2021.

⁴⁸ The United Cartel has accused one narcoblogger who operates on multiple social media platforms of favouring the Jalisco New Generation Cartel. “He works for the Jalisco New Generation Cartel. His job as a criminal is to misinform the population through social networks”. Post by alleged narcoblog, 27 March 2021. The narco-blog where this post appeared may itself be skewed toward the United Cartels, however. As an example of this former blogger’s alleged bias, to disprove claims that the CJNG used armed drones in an attack, this narcoblog showed videos that claimed to show members of the United Cartels flying drones. Facebook post, 14 April 2021.

reveal connections between the publisher and state security institutions, especially when those publications have access to confidential information not available to regular citizens.⁴⁹

B. *Winning Hearts and Minds*

Besides threatening rivals, criminal groups also turn to social media to win public support. Often groups claim to be protecting civilians from other, more dangerous outfits, or even the state, proclaiming they champion communities' "self-defence". For example, a Facebook post by a purported member of the United Cartels stated: "We don't want any more kidnappings in Michoacán. We're trying for a Michoacán free of CJNG, who dedicate themselves only to extortion, killing, etc. We want Michoacán to be for Michoacanos and that the townspeople feel safe".⁵⁰ The CJNG claimed the same of themselves: "we are here to serve the people".⁵¹ Groups will offer evidence of their purported noble aims and their enemies' brutality: after shootouts, for example, groups post content that is ostensibly from rivals' phones which they have snatched, including videos of torture.⁵²

Groups likewise distance themselves from activities they view as potentially unpopular. News outlets in Mexico widely reported that the Jalisco New Generation Cartel used drones armed with explosives in an attack on police in El Aguaje, Michoacán.⁵³ A post from a purported member countered: "Since we respect the authorities, we have to make it clear that United Cartels members are responsible for this act".⁵⁴ When rumours began circulating that the price of tortillas was rising due to the United Cartels increasing extortion payments, one of their pages responded that the hike was due simply to the market price of cornmeal: "United Cartels does not charge a fee to businesses. If the price goes up ... it is the maize".⁵⁵

Criminal rings also use social media to circulate information that paints them as protectors. Videos of criminal groups distributing aid (*narcodespensas*) to citizens in the wake of COVID-19 and natural disasters are now common.⁵⁶ For instance, the Jalisco Cartel distributed food and supplies during the pandemic in various cities as well as in poor rural areas in five states in Mexico, labelled the aid boxes with the

⁴⁹ Guadalupe Correa-Cabrera, Rajendra G. Kulkarni, Patrick R. Baxter and Naoru Koizumi, "Messengers of a Drug War in the Cyberspace: The Case of Tamaulipas", *Small Wars Journal*, 7 September 2021.

⁵⁰ Facebook post by an account purporting to be a member of Los Viagras, 10 January 2021.

⁵¹ Facebook post by an account claiming to be a CJNG member, 20 January 2021.

⁵² "When groups confront each other, and corpses remain, the winning hit men take cell phones, drugs, and their weapons. They unlock their phones and upload the content they find on the phones as a 'report'. Mainly videos of torture". Crisis Group telephone interview, narco-blog account, September 2021. An example of such a post reads: "We leave you with some images taken from the cell phone of a Jalisco after losing his cell phone in a confrontation, on it more proof of the government aiding the entry of the CJNG". Facebook post by group purporting to be the United Cartels, 21 January 2021.

⁵³ "Cártel Jalisco ataca con drones a personal de la SSP Michoacán", *El Universal*, 20 April 2021.

⁵⁴ Facebook post from purported CJNG account, 21 April 2021.

⁵⁵ Facebook post from purported United Cartels page, 10 May 2021.

⁵⁶ See, for example, the videos featured in "Narcodespensas: cuando el crimen "ayuda" a la gente", 5 May 2020; "En qué estados han repartido narcodespensas y por qué son estratégicos para el crimen organizado", *Infobae*, 29 April 2020.

group's name and posted them online – a clear effort at boosting their influence at a time of strain on official attention and resources.⁵⁷

These sorts of posts often fall into a grey zone for social media company policies. On one hand, platforms prohibit praise for criminal groups that feature on their lists of dangerous organisations, on the basis that it may glorify their actions.⁵⁸ YouTube, for example, removed a video from the early days of the pandemic showing the Jalisco Cartel distributing aid.⁵⁹ On the other hand, in the absence of clear-cut incitement to violence or graphic content, assessing whether these videos should be taken down is not always easy. Not all criminal groups are on lists of dangerous organisations. The content could be considered newsworthy, and it may not make any explicit reference to a particular criminal outfit. Moreover, while the assessment is being conducted, videos remain online where they can be picked up and circulated through mainstream press or narcoblogs.

C. *Narco-culture*

Criminal groups' most visible presence online comes in a somewhat different style: the promotion of *narcocultura*, or narco-culture. Typically, this content takes two forms. Narco-ballads (in Spanish, *narcocorridos*), which date to well before social media, praise the exploits of criminal leaders.⁶⁰ Such songs are frequently shared by group members on social media; they can even be commissioned on Facebook.⁶¹ Secondly, criminal group members share images and videos showcasing a luxurious, exciting lifestyle, particularly on primarily visual platforms such as Instagram and TikTok. "Cartel Tiktok" exploded in popularity in 2020, prompting one user – according to the *New York Times* – to ask, "Did the cartels just roll out their TikTok marketing strategy?"⁶² TikTok has since shadow-banned the hashtag #CartelTikTok, and is considering further measures such as improvements in computer vision to identify weapons or signs of cartel affiliation and removal of videos glorifying criminal lifestyles.⁶³

That said, TikTok cannot completely eradicate material exalting the cartels, because *narcocultura* exists outside and independent of cartel propaganda, as do other youth sub-cultures adopting *mafioso* style clothes or symbols.⁶⁴ Many who post about *narcocultura* are Mexican-Americans who have no direct experience of criminal groups

⁵⁷ "CJNG – Distributes Covid-19 food boxes in 6 states and Guatemala", *Borderland Beat*, 30 April 2020.

⁵⁸ Crisis Group telephone interview, TikTok employees, October 2022.

⁵⁹ "CJNG – Distributes Covid-19 food boxes in 6 states and Guatemala", *Borderland Beat*, 30 April 2020.

⁶⁰ Martín Meráz García, "Narcoballads': The Psychology and Recruitment Process of the 'Narc'", *Global Crime*, vol. 7, no. 2 (2006).

⁶¹ Facebook posts from narcocorrido groups, 21 and 27 April 2021.

⁶² Oscar López, "Guns, Drugs and Viral Content: Welcome to Cartel TikTok", *The New York Times*, 28 November 2020.

⁶³ Crisis Group telephone interview, Tik Tok employees, 28 October 2022.

⁶⁴ For example, Meta's Oversight Board reversed a decision by Instagram to remove a video in Britain featuring drill music, a subgenre of hip hop. The video and the song depicted images associated with gang violence and contained "veiled threats". The Oversight Board ultimately decided that the content was artistic in nature.

or links to them.⁶⁵ At the same time, TikTok would struggle to exclude content that might include certain elements of narco-culture but has no affiliation to criminal gangs or their members, especially given the size of the universe of users. Videos on TikTok mimicking the fashion and lifestyle associated with gang members or with the *buchonas*, the girlfriends of drug traffickers, are popular as far away as South East Asia.

Social media platforms struggle to draw a line between propaganda and subculture. Narco-ballads have a long history in Mexico, and even mainstream popular culture tends to glorify criminal groups.⁶⁶ One Facebook employee told Crisis Group that “promotion of narco-culture, especially through music, often toes the line when it comes to violating content [regulations]”.⁶⁷ A Twitter employee reiterated the same: “It’s prohibited to praise or support hate organisations, but in popular culture the way people talk about criminal entities can be pretty radical”.⁶⁸

D. Coordination

Criminal groups rarely coordinate their activities on public-facing platforms like Facebook and Twitter, for the obvious reason that their messages can be tracked. Lists of people illegal outfits planned to murder have circulated on Facebook, often from burner accounts, but in general criminal groups rely much more heavily on Encrypted Messaging Apps (EMAs), end-to-end encrypted platforms like WhatsApp and Messenger.⁶⁹ Encryption is particularly appealing for criminal groups because law enforcement cannot intercept the messages sent over those platforms. According to internal Meta research described to Crisis Group, criminals used Messenger to coordinate an attack on Mexico City’s police chief in 2020, among other means of communication.⁷⁰ The Sinaloa Cartel used WhatsApp to organise its violent attack on security forces in response to the attempted capture in 2019 of leader Joaquín “El Chapo” Guzmán’s son. Via the app, they offered money to anyone who joined the offensive, then called off the attack once El Chapo’s son was released.⁷¹ Members of organised crime also use WhatsApp and Messenger to extract extortion payments and to contact kidnapping victims’ families.⁷²

Platforms cannot moderate content sent through Encrypted Messaging Apps. For privacy reasons, nobody other than the sender and recipient can read messages sent

⁶⁵ Matt McAllester, “Mexico’s Narco Cultura: Glorifying Drug War Death and Destruction”, *Time*, 18 November 2013; and Carolina A. Miranda, “Influencer Jenny69 calls herself a ‘buchona.’ How a narco-inspired style came to rule social media”, *The Los Angeles Times*, 3 November 2021.

⁶⁶ “Narco-ballads are a difficult policy question”. Crisis Group telephone interview, Facebook employees, August 2021.

⁶⁷ Crisis Group telephone interview, Facebook employee, April 2021.

⁶⁸ Crisis Group telephone interview, Facebook employees, August 2021.

⁶⁹ Crisis Group interview with Facebook employee, April 2021, and Rebecca Plevin and Omar Ornelas, “‘We’re going to find you.’ Mexican cartels turn social media into tools for extortion, threats and violence”, *Desert Sun*, 27 February 2019.

⁷⁰ Crisis Group telephone interview, Facebook employee, April 2021.

⁷¹ Adyr Corarl and Victor Hugo Michel, “Narcos’ migran a WhatsApp para evadir servicios de inteligencia”, *El Milenio*, 27 January 2020. For more on the attack, see Falko Ernst, “Picking Up the Pieces after Mexico’s Criminal Siege”, Crisis Group Q&A, 22 October 2019.

⁷² Crisis Group telephone interview, Facebook employee, April 2021; Daniel Maza, “Aumentan los reports de extorsión vía Whatsapp”, *El Heraldo de Chiapas*, 5 February 2022.

by apps such as WhatsApp or Messenger. As a result, the platform managers simply do not know what content is being shared between users, even if these messages violate terms of use. Because of this, Meta employees told Crisis Group, they can only remove accounts that are used to make threats if the recipients of these messages report them.⁷³ Reporting, however, is not straightforward: when criminal groups contact a user, they often also threaten to inflict bodily harm, scaring people away from alerting the platform or authorities. The ease with which new accounts can be created may also induce a sense of hopelessness among victims, since the harasser can easily return under a different number. As a result, platform operators rarely see the most severe threats.

E. Recruitment

Recruitment is an area where both EMAs and public-facing platforms – particularly Facebook – have played a role. Online recruitment can occur in a number of ways. Vague, often high-paying job advertisements have been used to recruit smugglers and young men.⁷⁴ Users who purport to belong to criminal groups frequently post messages or images saying simply “we are recruiting ... Inbox”.⁷⁵ According to one Meta employee, however, the most common way that recruitment occurs on social media may be through individuals seeking out criminal affiliates’ accounts: “I heard anecdotally that at least half of the recruitment instances that are found are people approaching cartel posters saying, how do I join?”⁷⁶ In a recent Facebook whistleblower leak, the *Wall Street Journal* reported that internal research carried out by the social media company had confirmed that the Jalisco New Generation Cartel used the platform to recruit and train assassins, as well as pay them for services rendered. Unencrypted messages showed recruiters warning that, should they try to leave training, they would be seriously hurt or killed.⁷⁷

Public-facing social media can also be used to gather intelligence on rivals and plot armed manoeuvres. A high-ranking criminal operative from Michoacán reported to Crisis Group that the carelessness of young members of an enemy group had enabled him to plan an offensive against them: he ordered junior members of his group to comb Facebook to identify enemy operatives, pull down personal information such as photographs and locations, and assemble lists of possible targets. “You feel pity for them, they’re just kids posting all kinds of things ... but at the end of the day you have no choice since they [their bosses] send them to attack us”, he said.⁷⁸

⁷³ Crisis Group telephone interview, Facebook employee, April 2021.

⁷⁴ Crisis Group telephone interview, Facebook employee, April 2021; Julia Marnin, “How Cartels Use Facebook Ads to Trick Border Residents Into Smuggling Drugs into U.S.”, *Newsweek*, 22 June 2021; Antonio Rodríguez, “Expertos: Carteles reclutan a migrantes por redes sociales a cambio de ‘viaje seguro’ a EEUU”, *Telemundo 40*, 17 May 2021.

⁷⁵ For example, Facebook post by alleged criminal group member, 15 April 2021, accessed by Crisis Group 15 April 2021. See also, “El Cártel del Noreste recluta jóvenes en redes sociales por 18,000 pesos mensuales”, *Infobae*, 17 December 2021.

⁷⁶ Crisis Group telephone interview, Facebook employee, April 2021.

⁷⁷ Justin Scheck, Newley Purnell, and Jeff Horowitz, “Facebook Employees Flag Drug Cartels and Human Traffickers. The Company’s Response Is Weak, Documents Show”, *Wall Street Journal*, 16 September 2021.

⁷⁸ Crisis Group interview, Michoacán, April 2019.

IV. Challenges for Platforms

Social media platforms face a number of tough issues as they combat criminal groups' use of their services. Major social media companies track three groups that are priorities for law enforcement: terrorist organisations, hate groups, and criminal enterprises. Criminal groups are particularly hard to hone in on because – unlikely many terrorist and hate groups – they do not run forums devoted to discussing ideology and their propaganda is not designed to attract huge audiences.⁷⁹ As noted, they tend to operate through decentralised account networks, and rarely aim to go viral, which makes it more difficult to find offensive content.

The sheer multitude of criminal groups in Mexico – currently more than 200 – also makes it difficult for social media companies to keep up with all but a fraction of them. Platforms, including Twitter and Facebook, organise their responses to criminal groups in part on the basis of the official designation of particular entities as terrorist, hate, or criminal outfits made by the U.S. government. At Facebook, users and pages can also be taken down for consistently violating content standards, but not specifically for praising or depicting a criminal group – unless that group is on the “Dangerous Organizations” list maintained by Meta. Based in part on U.S. designations of violent organisations, groups can be added to that list following the company's internal deliberations.⁸⁰ It was not until 2020 that most of the largest outfits were specifically banned on Facebook, and to judge by a leak of documents from the company in 2021, only 21 Mexican criminal groups were banned and appeared on Meta's “Dangerous Organizations list”.⁸¹ Facebook employees explained to Crisis Group the challenges they encounter, observing that the way groups “split and splinter and fragment” makes it difficult to track them.⁸²

Further complicating the imposition of tighter controls by social media firms are *autodefensas*, or self-defence groups that have emerged at the local level, which tend to elude easy categorisation. Sometimes these groups serve as a form of community police, while in other cases they engage in criminal activity. Furthermore, criminal groups often adopt the language of self-defence as a cover for their activities. Social

⁷⁹ “Jihadists and neo-Nazis like to go into chat rooms and say crazy things – because they have some ideological bond. With cartels, there's no platforms or dark web forums where they go and chat with each other and discuss what they do. It's business”. Crisis Group telephone interview, Facebook employee, April 2021.

⁸⁰ “There has to be an internal case made for, say, banning the Sinaloa Cartel”, Crisis Group telephone interview, Facebook employee, April 2021; and Sam Biddle, “Revealed: Facebook's Secret Blacklist of ‘Dangerous Individuals and Organizations’”, *The Intercept*, 12 October 2021.

⁸¹ Crisis Group telephone interview, Facebook employee, April 2021. For more on delayed enforcement, see Justin Scheck, Newley Purnell, and Jeff Horowitz, “Facebook Employees Flag Drug Cartels and Human Traffickers. The Company's Response Is Weak, Documents Show”, *Wall Street Journal*, 16 September 2021. Crisis Group analysis of documents from Sam Biddle, see “Revealed: Facebook's Secret Blacklist of ‘Dangerous Individuals and Organizations’”, *The Intercept*, 12 October 2021.

⁸² Crisis Group telephone interview, Facebook employees, August 2021.

media platforms don't have a set policy for these self-defence forces, including when – or whether – they should be treated as violent actors themselves.⁸³

Efforts to curb the presence of criminal outfits on social media also run the risk of inadvertently eliminating information about dangerous hotspots. Telegram groups, Twitter pages, and Facebook pages – many of them narco-blogs – provide news about on-the-ground conditions that help keep civilians safe. This includes conveying information about the location of *narcobloqueos*, blockades set up to prevent entry or exit from certain territories; areas where conflict has broken out; and requests for help in finding kidnapping victims.⁸⁴

The line between helpful warnings and disinformation (ie, intentional efforts to deceive) can be extremely difficult to discern. Posts by criminal groups themselves can aid civilians in leaving areas that will in all likelihood be sites for a turf war between rival outfits.⁸⁵ They can, however, also sow confusion and stoke fear, particularly in the absence of more authoritative reporting. During an April 2021 escalation in Michoacán, a Facebook page affiliated with the United Cartels posted: “FOLLOWERS DON'T APPROACH THESE AREAS THE CJNG TERRORISTS ARE PICKING UP INNOCENT PEOPLE”.⁸⁶ A narco-blog that was seemingly affiliated with the Jalisco New Generation Cartel denied the United Cartels' claims, saying instead that the attack was a false-flag assault meant to discredit the group.⁸⁷

V. Reducing Harm and Protecting Information

The most direct route to reducing the real-world harm caused by criminal groups' online content is for platforms to invest more in detecting dangerous material such as recruitment campaigns, harassment and threats. Platforms generally have less capacity to do so outside the U.S. and Western Europe.⁸⁸ At Facebook, employees told Crisis Group that relative to hate and terrorist groups, platform responses to criminal groups are what they described as the “least mature”.⁸⁹ Mexican criminal

⁸³ “It's more difficult to draw the line with *autodefensas* versus a cartel”. Crisis Group telephone interview, Facebook employees, August 2021. “*Autodefensas* may not be covered”. Crisis Group telephone interview, Twitter employees, October 2021.

⁸⁴ For example, a Telegram post by Courage for Tamaulipas on 6 March 2023 featured a request for information about a missing woman, including name, description and photograph. One blogger reported: “I post about high-risk situations like shootings and chases; vehicles with armed people; blockades; safe houses and criminals from different drug trafficking organisations; I also provide alerts about the search for disappeared persons and social services”. Crisis Group interview, narco-blogger, September 2021.

⁸⁵ “Irrupción del CJNG en Tepalcatepec origina enfrentamientos y el levantamiento de armas de Cártels Unidos”, *Infobae*, 4 September 2021.

⁸⁶ Facebook post by purported United Cartels member, 29 April 2021.

⁸⁷ “The video that appears in *El Milenio* was recorded in the outskirts of the United Cartels' territory. ... Hitmen from Guicho with trucks with CJNG letters shot at passing vehicles so that people would report that they were Mencho's”. Facebook post by narco-blog allegedly tied to the CJNG, 30 April 2021.

⁸⁸ Jane Esberg, “What the Facebook Whistleblower Reveals about Social Media and Conflict”, Crisis Group Commentary, 18 November 2021.

⁸⁹ Crisis Group telephone interview, Facebook employees, August 2021.

groups' presence online is less easily identifiable than, for example, extremist jihadist groups or far right hate groups, which tend to operate within clear ideological parameters. This has meant that their posts frequently go undetected. According to one employee, for each post that is deleted for violating the rules ten others remain in place.⁹⁰

While it is admittedly hard to police the content criminal groups post, it is notable how much more effective social media platforms have been in eliminating posts by terrorist groups. To be sure, social media companies are dealing with more content given the sheer number of criminal groups and the decentralisation of their networks. Narcoculture makes this content ubiquitous, and countering it can raise concerns over curbs on freedom of expression. Moreover, as noted above, criminal groups generally lack the clearly defined ideological stances that make extremist groups easier to track.

Close cooperation with organisations that are well versed in local dynamics is essential to improving content moderation. Observatories on organised crime – which generally bring together members of civil society, academics and state officials to analyse criminal trends – could be particularly useful. Criminal groups have shown that they can adapt quickly to new content moderation rules, for example by using emojis to avoid shadow banning.⁹¹ By working closely with outside partners, social media platforms can better train both machine learning algorithms and human content moderators to keep up with these changes.⁹² Computer vision systems can be trained to automatically remove graphic content or symbols associated with specific cartels, while avoiding a blanket ban on everything related to narco-culture. Human content moderators should have an excellent grasp of the situation on the ground to evaluate cases falling in a grey area between culture and propaganda. It will not be possible to remove all narco-related content, but platforms can limit the ability of criminal groups to abuse social media.

Platforms should also be more sensitive to circumstances that can instigate rapid spikes in harmful social media content, including national elections and local escalations of conflict. This can help ensure that trusted on-the-ground voices are heard, and curb propaganda aimed at intimidating people or mobilising civilians on behalf of a criminal faction. More generally, social media platforms should make it easier for local researchers and civil society organisations to access their data. Network analysis and analysis of trends would allow platforms to remove clusters of accounts associated with known criminal group members, and reinforce the data available for content moderation adjustments. Prominent narcoblogs should be verified in coordination between social media platforms and trusted civil society partners, and companies should cooperate with these blogs to make sure that their content aligns with platform policies.

Content moderation, however, is being undermined by social media's own business practices, and platforms should contemplate changes to their algorithms to curb

⁹⁰ Crisis Group telephone interview, Facebook employee, April 2021.

⁹¹ Lieth Carrillo, "Chilean Gang Used Multiple Emojis on Instagram to Sell Drugs", *Organized Crime and Corruption Project*, 27 November 2023.

⁹² Ella Steen, Katherine Yurechko, and Daniel Klug, "You Can (Not) Say What You Want: Using Algorithmic Content Moderation on TikTok", *Social Media + Society*, vol. 9, no. 3 (July 2023).

use by criminal groups. Social media firms frequently gear their algorithms, or the rules governing what people see on social media, toward promoting user engagement – a practice that can play to criminal groups’ benefit.⁹³ Algorithms can promote content that includes harmful images or videos. In the course of our online research for this briefing, Facebook’s recommendation algorithm promoted accounts that prominently featured weapons and criminal insignia. Adjusting the ways in which recommendations are generated could help reduce the spread of criminal content and networks. Downgrading content that is largely image-based, and in its place promoting longer informative text, could encourage narco-bloggers to prioritise newsworthy pieces rather than amplify posts created or influenced by criminal groups.

This measure should go hand in hand with verification and protection programs, which remain the most powerful tool available to show platforms’ trust in accounts. Journalists on Facebook receive special security protections (Facebook Protect), which includes account monitoring for potential hacking.⁹⁴ Meta has increased its verification of civil society and journalist pages in conflict zones to promote trusted information.⁹⁵ Even so, no platform reports a specific policy outlining how they treat citizen journalists. Social media companies should remedy this in the case of Mexico by cultivating reliable regional and local sources who can help them to identify trusted anonymous information providers (that is, accounts that are widely viewed as relaying accurate, unbiased information), verify their accounts and amplify their voices.

Paid verification programs, on the other hand, could endanger attempts to limit false information. In March 2023, Elon Musk announced that Twitter would remove blue checkmarks from verified users, including journalists – a policy change that was put into effect, with some exceptions, a month later. A paid plan (“Twitter Blue”) now lets any user receive a blue checkmark next to their profile, making it far more difficult to verify whether a source can be trusted.

Finally, outside the digital realm, senior government officials need to allocate more energy and resources to protecting professional journalists, and to rein in verbal attacks on the press that could be seen as giving license for physical abuses. Criminal groups have thrived online in part by purporting to provide on-the-ground information in news deserts. The government of President López Obrador expanded a government program called the Federal Mechanism for the Protection of Human Rights Defenders and Journalists, designed to provide security to journalists and human rights defenders by offering an array of protective measures, from camera systems and panic buttons to police protection or relocation for those under threat.⁹⁶ Although the Mechanism has reportedly protected the lives of a number of journalists, it has not ended the violence they face, with a number of journalists killed while

⁹³ Katherine J. Wu, “Radical ideas spread through social media. Are the algorithms to blame?”, PBS, 28 March 2019; Areeba Shah, “TikTok’s algorithm is pushing out extremist and violent content to 13-year-olds”, *Salon*, 25 March 2023.

⁹⁴ See “Facebook Protect”.

⁹⁵ Crisis Group Africa Report N°295, *Easing Cameroon’s Ethno-political Tensions, On- and Off-line*, 3 December 2020.

⁹⁶ “Conoce más sobre el Mecanismo de protección de personas defensoras de derechos humanos y periodistas”, Mexican government, November 2018

under protection and others murdered while waiting for their application to be approved. Impunity for the killings of journalists remains high.⁹⁷

The Mexican president has meanwhile contributed to insecurity through hostility toward journalists. López Obrador has attacked the press on various occasions as “mercenaries”, “thugs” and “sellouts”, comments that journalists believe have intensified the dangers they face – a practice that should stop.⁹⁸ Meanwhile, protection schemes for journalists are understaffed and under-resourced, and there are long delays between a journalist’s request for protection and its eventual supply. Federal forces have also found it difficult to coordinate with local and state police to provide protection on the ground. Better staffing, clearer pathways of communication, and expansion of the program so that more journalists can access it in a timely fashion all stand high among the improvements needed.⁹⁹

Foreign donors should also sponsor the creation of secure, encrypted, online collaborative platforms for reporters to deposit information and share their investigative work in the event of their death, which might help deter intimidation and violence against individual journalists.¹⁰⁰

VI. Conclusion

Illegal outfits in Mexico use digital platforms to spread propaganda, recruit new members and make threats against rivals. Too often they are several steps ahead of those who would regulate their on-line activities. Decentralisation, the use of multiple accounts, and the ability to work through members and supporters make it hard to track criminal groups’ posts. Adding to the confusion, members of the general public frequently put their own narco-inspired content on line, complicating platforms’ efforts to crack down on content glorifying crime. With professional journalists unable to access the most dangerous parts of the country, an informal array of criminal groups, narco-bloggers and citizen journalists fill the gap with content that may or may not be reliable.

Mindful of the need to protect free expression, there is more that both social media companies and government officials can do to make the digital realm a less welcoming place for criminal groups, and to curb their ability to advance their interests on line. Platforms should work more closely with local civil society to identify criminal online actors and better understand the content they are putting up, so that moderation efforts can be more effectively tailored. Certain algorithm adjustments

⁹⁷ Samedí Aguirre, “Violencia contra periodistas en México: más de 90% de crímenes impunes y mecanismos de protección insuficientes”, *Animal Político*, 25 January 2022.

⁹⁸ Carmen de Carlos, “La prensa mexicana exige a López Obrador que cese con su “hostigamiento” para evitar más muertes”, *El Debate*, 22 December 2022; Javier Villaverde, “López Obrador, cuestionado por la prensa mexicana por los múltiples asesinatos a periodistas”, *La Razón*, 10 May 2022; “As journalists face deadly violence, Mexico’s president attacks the media” (editorial), *Washington Post*, 15 February 2022.

⁹⁹ Jan-Albert Hootsen, “López Obrador’s last chance to protect the press”, *Just Security*, 23 March 2022.

¹⁰⁰ See for example the case of the late Colombia journalist Rafael Moreno. “Assassination of Rafael Moreno: Revealing the Missions Embezzled in Colombia”, *Forbidden Stories*, 18 April 2023.

can also help prioritise quality informative content over material that glorifies crime. Finally, back in the physical world, the government can and should do more to protect professional journalists through stronger and more agile interventions, and by sending a consistent signal, at every level of government, that a free and vibrant press is essential to the well-being of the nation and all who live there.

Mexico City/New York/Bogotá/Brussels, 31 January 2024



International Crisis Group

Headquarters

Avenue Louise 235, 1050 Brussels, Belgium

Tel: +32 2 502 90 38

brussels@crisisgroup.org

New York Office

newyork@crisisgroup.org

Washington Office

washington@crisisgroup.org

London Office

london@crisisgroup.org

Regional Offices and Field Representation

Crisis Group also operates out of over 25 locations in Africa, Asia, Europe, the Middle East and Latin America.

See www.crisisgroup.org for details

PREVENTING WAR. SHAPING PEACE.