

# Revealed: Facebook's Secret Blacklist Of "Dangerous Individuals And Organizations"

Tuesday 19 October 2021, by [BIDDLE Sam](#) (Date first published: 13 October 2021).

**Experts say the public deserves to see the list, a clear embodiment of U.S. foreign policy priorities that could disproportionately censor marginalized groups.**

To ward off accusations that it helps terrorists spread propaganda, Facebook has for many years barred users from speaking freely about people and groups it says promote violence.

The restrictions appear to trace back to 2012, when in the face of growing alarm [in Congress](#) and [the United Nations](#) about online terrorist recruiting, Facebook [added](#) to its Community Standards a ban on "organizations with a record of terrorist or violent criminal activity." This modest rule has since ballooned into what's known as the Dangerous Individuals and Organizations policy, a sweeping set of restrictions on what Facebook's nearly 3 billion users can say about an enormous and ever-growing roster of entities deemed beyond the pale.

In recent years, the policy has been used at a more rapid clip, including against the president of the United States, and taken on almost totemic power at the social network, trotted out to reassure the public whenever paroxysms of violence, from [genocide in Myanmar](#) to [riots on Capitol Hill](#), are linked to Facebook. Most recently, following a damning series of Wall Street Journal articles showing the company [knew](#) it facilitated [myriad offline harms](#), a Facebook vice president cited the policy as evidence of the company's diligence in an internal memo [obtained](#) by the New York Times.

But as with other attempts to limit personal freedoms in the name of counterterrorism, Facebook's DIO policy has become an unaccountable system that disproportionately punishes certain communities, critics say. It is built atop a blacklist of over 4,000 people and groups, including politicians, writers, charities, hospitals, hundreds of music acts, and long-dead historical figures.

A range of legal scholars and civil libertarians have called on the company to publish the list so that users know when they are in danger of having a post deleted or their account suspended for praising someone on it. The company has repeatedly refused to do so, claiming it would endanger employees and permit banned entities to circumvent the policy. Facebook did not provide The Intercept with information about any specific threat to its staff.

Despite Facebook's claims that disclosing the list would endanger its employees, the company's hand-picked Oversight Board has formally [recommended publishing all of it on multiple occasions](#), as recently as August, because the information is in the public interest.

The Intercept has reviewed a snapshot of the full DIO list and is today publishing a [reproduction of the material in its entirety](#), with only minor redactions and edits to improve clarity. It is also publishing an [associated policy document](#), created to help moderators decide what posts to delete and what users to punish.

"Facebook puts users in a near-impossible position by telling them they can't post about dangerous

groups and individuals, but then refusing to publicly identify who it considers dangerous,” said Faiza Patel, co-director of the Brennan Center for Justice’s liberty and national security program, who reviewed the material.

The list and associated rules appear to be a clear embodiment of American anxieties, political concerns, and foreign policy values since 9/11, experts said, even though the DIO policy is meant to protect all Facebook users and applies to those who reside outside of the United States (the vast majority). Nearly everyone and everything on the list is considered a foe or threat by America or its allies: Over half of it consists of alleged foreign terrorists, free discussion of which is subject to Facebook’s harshest censorship.

The DIO policy and blacklist also place far looser prohibitions on commentary about predominately white anti-government militias than on groups and individuals listed as terrorists, who are predominately Middle Eastern, South Asian, and Muslim, or those said to be part of violent criminal enterprises, who are predominantly Black and Latino, the experts said.

The materials show Facebook offers “an iron fist for some communities and more of a measured hand for others,” said Ángel Díaz, a lecturer at the UCLA School of Law who has researched and [written on the impact](#) of Facebook’s moderation policies on marginalized communities.

Facebook’s policy director for counterterrorism and dangerous organizations, Brian Fishman, said in a written statement that the company keeps the list secret because “[t]his is an adversarial space, so we try to be as transparent as possible, while also prioritizing security, limiting legal risks and preventing opportunities for groups to get around our rules.” He added, “We don’t want terrorists, hate groups or criminal organizations on our platform, which is why we ban them and remove content that praises, represents or supports them. A team of more than 350 specialists at Facebook is focused on stopping these organizations and assessing emerging threats. We currently ban thousands of organizations, including over 250 white supremacist groups at the highest tiers of our policies, and we regularly update our policies and organizations who qualify to be banned.”

Though the experts who reviewed the material say Facebook’s policy is unduly obscured from and punitive toward users, it is nonetheless a reflection of a genuine dilemma facing the company. After the Myanmar genocide, the company recognized it had become perhaps the most powerful system ever assembled for the global algorithmic distribution of violent incitement. To do nothing in the face of this reality would be viewed as grossly negligent by vast portions of the public — even as Facebook’s attempts to control the speech of billions of internet users around the world is widely seen as the stuff of autocracy. The DIO list represents an attempt by a company with a historically unprecedented concentration of power over global speech to thread this needle.

### **Harsher Restrictions for Marginalized and Vulnerable Populations**

The list, the foundation of Facebook’s Dangerous Individuals and Organizations policy, is in many ways what the company has described in the past: a collection of groups and leaders who have threatened or engaged in bloodshed. The snapshot reviewed by The Intercept is separated into the categories Hate, Crime, Terrorism, Militarized Social Movements, and Violent Non-State Actors. These categories were organized into a system of three tiers under [rules](#) rolled out by Facebook in late June, with each tier corresponding to speech restrictions of varying severity.

But while labels like “terrorist” and “criminal” are conceptually broad, they look more like narrow racial and religious proxies once you see how they are applied to people and groups in the list, experts said, raising the likelihood that Facebook is placing discriminatory limitations on speech.

Regardless of tier, no one on the DIO list is allowed to maintain a presence on Facebook platforms, nor are users allowed to represent themselves as members of any listed groups. The tiers determine instead what other Facebook users are allowed to say about the banned entities. Tier 1 is the most strictly limited; users may not express anything deemed to be praise or support about groups and people in this tier, even for nonviolent activities (as determined by Facebook). Tier 1 includes alleged terror, hate, and criminal groups and alleged members, with terror defined as “organizing or advocating for violence against civilians” and hate as “repeatedly dehumanizing or advocating for harm against” people with protected characteristics. Tier 1’s criminal category is almost entirely American street gangs and Latin American drug cartels, predominantly Black and Latino. Facebook’s terrorist category, which is 70 percent of Tier 1, overwhelmingly consists of Middle Eastern and South Asian organizations and individuals — who are disproportionately represented throughout the DIO list, across all tiers, where close to 80 percent of individuals listed are labeled terrorists.

Facebook takes most of the names in the terrorism category directly from the U.S. government: Nearly 1,000 of the entries in the dangerous terrorism list note a “designation source” of “SDGT,” or Specially Designated Global Terrorists, a sanctions list maintained by the Treasury Department and created by George W. Bush in the immediate aftermath of the September 11 attacks. In many instances, names on Facebook’s list include passport and phone numbers found on the official SDGT list, suggesting entries are directly copied over.

Other sources cited include the Terrorism Research & Analysis Consortium, a private subscription-based database of purported violent extremists, and SITE, a private terror-tracking operation with a [long, controversial history](#). “An Arabic word can have four or five different meanings in translation,” Michael Scheuer, the former head of the CIA’s Osama bin Laden unit, [told](#) the New Yorker in 2006, noting that he thinks SITE typically chooses the “most warlike translation.” It appears Facebook has worked with its tech giant competitors to compile the DIO list; one entry carried a note that it had been “escalated by” a high-ranking staffer at Google who previously worked in the executive branch on issues related to terrorism. (Facebook said it does not collaborate with other tech companies on its lists.)

There are close to 500 hate groups in Tier 1, including the more than 250 white supremacist organizations Fishman referenced, but Faiza Patel, of the Brennan Center, noted that hundreds of predominantly white right-wing militia groups that seem similar to the hate groups are “treated with a light touch” and placed in Tier 3.

Tier 2, “Violent Non-State Actors,” consists mostly of groups like armed rebels who engage in violence targeting governments rather than civilians, and includes many factions fighting in the Syrian civil war. Users can praise groups in this tier for their nonviolent actions but may not express any “substantive support” for the groups themselves.

Tier 3 is for groups that are not violent but repeatedly engage in hate speech, seem poised to become violent soon, or repeatedly violate the DIO policies themselves. Facebook users are free to discuss Tier 3 listees as they please. Tier 3 includes Militarized Social Movements, which, judging from its DIO entries, is mostly right-wing American anti-government militias, which are virtually entirely white.

“The lists seem to create two disparate systems, with the heaviest penalties applied to heavily Muslim regions and communities,” Patel wrote in an email to The Intercept. The differences in demographic composition between Tiers 1 and 3 “suggests that Facebook — like the U.S. government — considers Muslims to be the most dangerous.” By contrast, Patel pointed out, “Hate groups designated as Anti-Muslim hate groups by the [Southern Poverty Law Center](#) are

overwhelmingly absent from Facebook's lists."

Anti-government militias, among those receiving more measured interventions from Facebook, "present the most lethal [domestic violent extremist] threat" to the U.S., intelligence officials [concluded](#) earlier this year, a view shared by many nongovernmental researchers. A crucial difference between alleged foreign terror groups and say, the Oath Keepers, is that domestic militia groups have considerable political capital and support on the American right. The Militarized Social Movement entries "do seem to be created in response to more powerful organizations and ethnic groups breaking the rules pretty regularly," said Ángel Díaz, of UCLA School of Law, "and [Facebook] feeling that there needs to be a response, but they didn't want the response to be as broad as it was for the terrorism portion, so they created a subcategory to limit the impact on discourse from politically powerful groups." For example, the extreme-right movement known as "boogaloo," which [advocates for a second Civil War](#), is considered a Militarized Social Movement, which would make it subject to the relatively lenient Tier 3 rules. Facebook has only classified as Tier 1 a subset of boogaloo, which it made clear was "[distinct](#) from the broader and loosely-affiliated boogaloo movement."

A Facebook spokesperson categorically denied that Facebook gives extremist right-wing groups in the U.S. special treatment due to their association with mainstream conservative politics. They added that the company tiers groups based on their behavior, stating, "Where American groups satisfy our definition of a terrorist group, they are designated as terrorist organizations (E.g. The Base, Atomwaffen Division, National Socialist Order). Where they satisfy our definition of hate groups, they are designated as hate organizations (For example, Proud Boys, Rise Above Movement, Patriot Front)."

The spokesperson framed the company's treatment of militias as one of aggressive regulation rather than looseness, saying Facebook's list of 900 such groups "is among the the most robust" in the world: "The Militarized Social Movement category was developed in 2020 explicitly to expand the range of organizations subject to our DOI policies precisely because of the changing threat environment. Our policy regarding militias is the strongest in the industry."

On the issue of how Facebook's tiers often seem to sort along racial and religious lines, the spokesperson cited the presence of the white supremacists and hate groups in Tier 1 and said "focusing solely on" terrorist groups in Tier 1 "is misleading." They added: "It's worth noting that our approach to white supremacist hate groups and terrorist organization is far more aggressive than any government's. All told, the United Nations, European Union, United States, United Kingdom, Canada, Australia, and France only designate thirteen distinct white supremacist organizations. Our definition of terrorism is public, detailed and was developed with significant input from outside experts and academics. Unlike some other definitions of terrorism, our definition is agnostic to religion, region, political outlook, or ideology. We have designated many organizations based outside the Middle Eastern and South Asian markets as terrorism, including orgs based in North America and Western Europe (including the National Socialist Order, the Feurerkrieg Division, the Irish Republican Army, and the National Action Group)."

On Facebook's list, however, the number of listed terrorist groups based in North American or Western Europe amounts to only a few dozen out of over a thousand.

Though the list includes a litany of ISIS commanders and Al Qaeda militants whose danger to others is uncontroversial, it would be difficult to argue that some entries constitute much of a threat to anyone at all. Due to the company's mimicry of federal terror sanctions, which are meant to punish international adversaries rather than determine "dangerousness," it is Facebook policy that the likes of the Iran Tractor Manufacturing Company and the [Palestinian Relief and Development Fund](#), a

U.K.-based aid organization, are both deemed too much of a real-world danger for free discussion on Facebook and are filed among Tier 1 terrorist organizations like al-Shabab.

“When a major, global platform chooses to align its policies with the United States — a country that has long exercised hegemony over much of the world (and particularly, over the past twenty years, over many predominantly Muslim countries), it is simply recreating those same power differentials and taking away the agency of already-vulnerable groups and individuals,” said Jillian York, director for international freedom of expression at the Electronic Frontier Foundation, who also reviewed the reproduced Facebook documents.

Facebook’s list represents an expansive definition of “dangerous” throughout. It includes the deceased 14-year-old Kashmiri child soldier Mudassir Rashid Parray, over 200 musical acts, television stations, a video game studio, airlines, [the medical university](#) working on Iran’s homegrown Covid-19 vaccine, and many long-deceased historical figures like [Joseph Goebbels](#) and Benito Mussolini. Including such figures is “fraught with problems,” a group of University of Utah social media researchers [recently told](#) Facebook’s Oversight Board.

### **Troubling Guidelines for Enforcement**

Internal Facebook materials walk moderators through the process of censoring speech about the blacklisted people and groups. The materials, portions of which were [previously reported by The Guardian](#) and [Vice](#), attempt to define what it means for a user to “praise,” “support,” or “represent” a DIO listee and detail how to identify prohibited comments.

Although Facebook [provides a public set of such guidelines](#), it publishes only limited examples of what these terms mean, rather than definitions. Internally, it offers not only the definitions, but also much more detailed examples, including a dizzying list of hypotheticals and edge cases to help determine what to do with a flagged piece of content.

Facebook’s global content moderation workforce, an [outsourced army](#) of [hourly contractors](#) frequently [traumatized](#) by the graphic nature of their work, are expected to use these definitions and examples to figure out if a given post constitutes forbidden “praise” or meets the threshold of “support,” among other criteria, shoehorning the speech of billions of people from hundreds of countries and countless cultures into a tidy framework decreed from Silicon Valley. Though these workers operate in tandem with automated software systems, determining what’s “praise” and what isn’t frequently comes down to personal judgment calls, assessing posters’ intent. “Once again, it leaves the real hard work of trying to make Facebook safe to outsourced, underpaid and overworked content moderators who are forced to pick up the pieces and do their best to make it work in their specific geographic location, language and context,” said Martha Dark, the director of Foxglove, a legal aid group that works with moderators.

In the internal materials, Facebook essentially says that users are allowed to speak of Tier 1 entities so long as this speech is neutral or critical, as any commentary considered positive could be construed as “praise.” Facebook users are barred from doing anything that “seeks to make others think more positively” or “legitimize” a Tier 1 dangerous person or group or to “align oneself” with their cause — all forms of speech considered “praise.” The materials say, “Statements presented in the form of a fact about the entity’s motives” are acceptable, but anything that “glorifies the entity through the use of approving adjectives, phrases, imagery, etc” is not. Users are allowed to say that a person Facebook considers dangerous “is not a threat, relevant, or worthy of attention,” but they may not say they “stand behind” a person on the list they believe was wrongly included — that’s considered aligning themselves with the listee. Facebook’s moderators are similarly left to decide for themselves what constitutes dangerous “glorification” versus permitted “neutral speech,” or what



counts as “academic debate” and “informative, educational discourse” for billions of people.

Determining what content meets Facebook’s definitions of banned speech under the policy is a “struggle,” according to a Facebook moderator working outside of the U.S. who responded to questions from The Intercept on the condition of anonymity. This person said analysts “typically struggle to recognize political speech and condemnation, which are permissible context for DOI.” They also noted the policy’s tendency to misfire: “[T]he fictional representations of [dangerous individuals] are not allowed unless shared in a condemning or informational context, which means that sharing a Taika Waititi [photo](#) from [the film] Jojo Rabbit will get you banned, as well as a [meme](#) with the actor playing Pablo Escobar (the one in the empty swimming pool).”

These challenges are compounded because a moderator must try to gauge how their fellow moderators would assess the post, since their decisions are compared. “An analyst must try to predict what decision would a quality reviewer or a majority of moderators take, which is often not that easy,” the moderator said.

The rules are “a serious risk to political debate and free expression,” Patel said, particularly in the Muslim world, where DIO-listed groups exist not simply as military foes but part of the sociopolitical fabric. What looks like glorification from a desk in the U.S. “in a certain context, could be seen [as] simple statements of facts,” EFF’s York agreed. “People living in locales where so-called terrorist groups play a role in governance need to be able to discuss those groups with nuance, and Facebook’s policy doesn’t allow for that.”

As Patel put it, “A commentator on television could praise the Taliban’s promise of an inclusive government in Afghanistan, but not on Facebook.”

The moderator working outside of the U.S. agreed that the list reflects an Americanized conception of danger: “The designations seem to be based on American interests,” which “does not represent the political reality in those countries” elsewhere in the world, the person said.

Particularly confusing and censorious is Facebook’s definition of a “Group Supporting Violent Acts Amid Protests,” a subcategory of Militarized Social Movements barred from using the company’s platforms. Facebook describes such a group as “a non-state actor” that engages in “representing [or] depicting ... acts of street violence against civilians or law enforcement,” as well as “arson, looting, or other destruction of private or public property.” As written, this policy would appear to give Facebook license to apply this label to virtually any news organization covering — that is to say, depicting — a street protest that results in property damage, or to punish any participant uploading pictures of these acts by others. Given the praise piled onto Facebook a decade ago for the belief it had helped drive the Arab Spring uprisings across North Africa and the Middle East, it’s notable that, say, an Egyptian organization documenting violence amid the protests in Tahrir Square in 2011 could be deemed a dangerous Militarized Social Movement under 2021’s rulebook.

Díaz, of UCLA, told The Intercept that Facebook should disclose far more about how it applies these protest-related rules. Will the company immediately shut down protest organizing pages the second any fires or other property damage occurs? “The standards that they’re articulating here suggest that [the DIO list] could swallow up a lot of active protesters,” Díaz said.

It’s possible protest coverage was linked to the DIO listing of [two anti-capitalist media organizations](#): Crimethinc and It’s Going Down. Facebook banned both publications in 2020, citing DIO policy, and both are indeed found on the list, designated as Militarized Social Movements and further tagged as “armed militias.”

A representative for It's Going Down, who requested anonymity on the basis of their safety, told The Intercept that "outlets across the political spectrum report on street clashes, strikes, riots, and property destruction, but here Facebook seems to be imply if they don't like what analysis ... or opinion one writes about why millions of people took to the streets last summer during the pandemic in the largest outpouring in U.S. history, then they will simply remove you from the conversation." They specifically denied that the group is an armed militia, or even activist or a social movement, explaining that it is instead a media platform "featuring news, opinion, analysis and podcasts from an anarchist perspective." A representative of Crimethinc likewise denied that the group is armed or "'militarized' in any sense. It is a news outlet and book publisher, like Verso or Jacobin." The representative requested anonymity citing right-wing threats to the organization.

Facebook did not address questions about why these media organizations had been internally designated "armed militias" but instead, when asked about them, reiterated its prohibition on such groups and on Groups Supporting Violent Acts Amid Protests.

Facebook's internal moderation guidelines also leave some puzzling loopholes. After the platform [played a role](#) in facilitating a genocide in Myanmar, company executive Alex Warofka wrote, "We agree that we can and should do more" to "prevent our platform from being used to foment division and incite offline violence." But Facebook's ban against violent incitement is relative, expressly permitting, in the policy materials obtained by The Intercept, calls for violence against "locations no smaller than a village." For example, cited as fair game in the rules is the statement "We should invade Libya." The Facebook spokesperson, said, "The purpose of this provision is to allow debate about military strategy and war, which is a reality of the world we live in," and acknowledged that it would allow for calls of violence against a country, city, or terrorist group, giving as an example of a permitted post under the last category a statement targeting an individual: "We should kill Osama bin Laden."

## **Harsh Suppression of Speech About the Middle East**

Enforcing the DIO rules leads to some surprising outcomes for a company that [claims "free expression" as a core principle](#). In 2019, [citing the DIO policy](#), Facebook blocked an online university symposium [featuring Leila Khaled](#), who participated in two plane hijackings in the 1960s in which no passengers were hurt. Khaled, now 77, is still present in the version of Facebook's terrorism list obtained by The Intercept. In February, Facebook's internal Oversight Board moved to reverse a decision to delete a post questioning the imprisonment of leftist Kurdish revolutionary Abdullah Öcalan, a DIO listee whom the U.S. [helped Turkish intelligence forces abduct](#) in 1999.

In July, journalist Rania Khalek [posted a photo to Instagram](#) of a billboard outside Baghdad International Airport depicting Iranian general Qassim Suleimani and Iraqi military commander Abu Mahdi al-Muhandis, both assassinated by the United States and both on the DIO list. Khalek's Instagram upload was quickly deleted for violating what a notification called the "violence or dangerous organizations" policy. In an email, Khalek told The Intercept, "My intent when I posted the photo was to show my surroundings," and "the fact that [the billboard is] so prominently displayed at the airport where they were murdered shows how they are perceived even by Iraqi officialdom."

More recently, Facebook's DIO policy collided with the Taliban's toppling of the U.S.-backed government in Afghanistan. After the Taliban assumed control of the country, Facebook announced the group was [banned](#) from having a presence on its apps. Facebook now finds itself in the position of not just censoring an entire country's political leadership but placing serious constraints on the public's ability to discuss or even merely depict it.

Other incidents indicate that the DIO list may be too blunt an instrument to be used effectively by Facebook moderators. In May, Facebook deleted a variety of posts by Palestinians attempting to document Israeli state violence at Al Aqsa Mosque, the third holiest site in Islam, because company staff mistook it for an unrelated organization on the DIO list with “Al-Aqsa” in its name (of which there are several), judging from an internal memo [obtained](#) by BuzzFeed News. Last month, Facebook censored an Egyptian user who posted an Al Jazeera article about the Al-Qassam Brigades, a group active in neighboring Palestine, along with a caption that read simply “Ooh” in Arabic. Al-Qassam does not appear on the DIO list, and Facebook’s Oversight Board [wrote](#) that “Facebook was unable to explain why two human reviewers originally judged the content to violate this policy.”

While the past two decades have inured many the world over to secret ledgers and laws like [watchlists](#) and no-fly bans, Facebook’s privatized version indicates to York that “we’ve reached a point where Facebook isn’t just abiding by or replicating U.S. policies, but going well beyond them.”

“We should never forget that nobody elected Mark Zuckerberg, a man who has never held a job other than CEO of Facebook.”

---

**Sam Biddle**

[Click here](#) to subscribe to ESSF newsletters in English and/or French.

---

**P.S.**

The Intercept

<https://theintercept.com/2021/10/12/facebook-secret-blacklist-dangerous/>