

“Shaheed” and designated dangerous individuals (PAO 2023-01)

March 9, 2023

Submit public comment [here](#).

The Oversight Board has accepted Meta’s request for a policy advisory opinion on its approach to moderating the Arabic word “shaheed” when referring to individuals it classifies as “dangerous,” including terrorists. The word has multiple meanings but is often translated as “martyr,” and accounts for more content removals under the Community Standards than any other single word or phrase on Meta’s platforms. The company acknowledges that its current approach may result in significant over-enforcement, particularly in Arabic-speaking countries, and has explored alternatives. However, it points to the difficulties and tensions in moderating use of the term at scale.

Meta’s request to the Board is available in [English](#) and [Arabic](#).

Meta estimates that the word “shaheed,” and its variations, account for more content removals under the Community Standards than any other single word or phrase on its platforms.

In its request, Meta asks the Board whether it should continue to remove content using “shaheed” to refer to individuals designated as dangerous under its Dangerous Individuals and Organizations policy, or whether a different approach would better align with the company’s values and human rights responsibilities. Meta also requests guidance on similar content issues that may arise in the future.

Meta says it removes content referring to designated dangerous individuals as “shaheed” because it translates the word as “martyr.” It therefore considers it a form of praise. Praising a designated individual is prohibited under the Dangerous Individuals and Organizations policy. However, the company acknowledges that the meaning of “shaheed” varies.

In its request, Meta describes the word “shaheed” as an “honorific” term, used by many communities around the world, across cultures, religions, and languages. The company says the term has “multiple meanings” and is “used to describe someone dying unexpectedly or prematurely, at times referring to an honourable death, such as when one dies in an accident or in a conflict or war.” Meta states that the common English translation is “martyr,” and assumes this meaning for the purposes of content moderation, in all contexts. However, it notes that “there is no direct equivalent to the term in the English language.”

The Dangerous Individuals and Organizations policy prohibits “praise, substantive support, or representation of designated entities and individuals.” Its definition of praise includes, giving “a designated entity or event a sense of achievement,” legitimizing “the cause of a designated entity,” and aligning “oneself ideologically with a designated entity or event.” This definition was added following a recommendation by the Board (“Nazi quote” case recommendation two). Because Meta assumes “shaheed” means “martyr,” it is considered a form of praise when used to refer to a designated entity. The Board previously recommended that Meta publish its list of designated entities, or illustrative examples (“Nazi quote” case, recommendation three). Meta has not published the list and provided no further updates on this recommendation following a feasibility assessment.

Removal of the word “shaheed” can result in severe “strikes,” or sanctions, for users. The company acknowledges that its current approach may result in significant over-enforcement, particularly in Arabic-speaking countries. Given the multiple meanings of “shaheed” and difficulties in accounting for context at scale, Meta accepts that it may be removing speech that is “not intended to praise a designated individual.” For example, where “shaheed” is used to refer to a premature death or a deceased person, rather than to glorify their conduct. Meta does not apply its policy exception for neutral news reporting to the word “shaheed,” as it assumes the word not to be neutral.

Because of these concerns, Meta initiated a policy development process in 2020 to reassess its use of the term “shaheed.” This included a research review and stakeholder consultation. Meta describes as key findings of this stakeholder engagement that the meaning of “shaheed” depends on context, and that in some instances the term has become desensitized and disconnected from praise. During this process, Meta identified two scalable policy options for use of the word “shaheed.” However, each had drawbacks, there was no consensus among stakeholders, and Meta did not settle on a new approach. The company emphasizes that due to the volume of content on its platform, a key practical concern is whether enforcement works at scale.

Questions posed by Meta to the Board:

Meta presented the following policy options to the Board for its consideration:

- 1. Continue its current approach and remove content that uses “shaheed” to refer to an individual designated as dangerous under the Dangerous Individuals and Organizations policy. Meta recognizes that this option can restrict users’ voice and leads to the greatest amount of content being removed. The company also recognizes that this approach can result in severe “strikes” for users, which can lead to their account being disabled. It can also disproportionately affect “specific communities for whom “shaheed” is a common word,” raising concerns about equality and fairness. However, Meta says this option pursues its value of “Safety” and is the easiest to operate.*

2. Allow content that refers to a designated individual as “shaheed” when the following conditions are met: (i) it is used in a context that is permitted under the Dangerous Individuals and Organizations policy (e.g., condemnation, news reporting, academic debate); (ii) there is no additional praise, representation or support of a designated individual (e.g., the post does not explicitly praise a perpetrator of a terrorist attack or legitimize their violence); and (iii) there is no signal of violence in the content (e.g., a visual depiction of weapons, military language, or references to real-world violence). Meta says this would support its value of “Voice,” though it may impact its value of “Safety,” and would better align with the Dangerous Individuals and Organizations policy, which allows discussion of designated individuals. This option would also reduce negative impacts on communities where “shaheed” is commonly used, and “may better enable news outlets ...to provide objective news coverage.” Meta notes, however, that targets of terrorism and violence may object to changing the policy in this way. It also says this option would make enforcement more complex, which could lead to enforcement errors.

3. Remove content that uses “shaheed” to refer to an individual designated as dangerous under Meta’s Dangerous Individuals and Organizations policy only where there is additional praise, representation or support, or where there is a signal of violence. Meta believes that this option would “better align with Meta’s value of voice and principles of international law” but “could be perceived as promoting voice over the value of safety.” This option “maximizes the way Shaheed could be used,” allowing people to “use the word according to their respective culture or vernacular.” However, it could lead to content on its platforms that intends to legitimize terrorism. Meta says this option would be easier to operate than option two.

While the Board will consider the specific options provided by Meta, the Board's recommendations and policy advisory opinion might not be limited to these options.

The Board requests public comments that address:

- Examples of how Meta’s current approach to “shaheed” as praise impacts freedom of expression on Instagram and Facebook, especially for civil society, journalists, and human rights defenders in regions where the word is commonly used.
- Research into the connection between restricting praise of individuals associated with terrorist organizations on social media and the effective prevention of terrorist acts.
- How Meta should account for the variety of meanings and diverse cultural contexts for using the term “shaheed” in different regions, languages and dialects, given the trade-offs inherent in enforcing content policies at scale, and the implications for Meta’s responsibility to respect human rights.
- What processes and safeguards should be in place to mitigate the risks of under- or over-enforcement of the Dangerous Individuals and Organizations policy, in particular across diverse cultures, languages and dialects.

- *How to measure the accuracy of policy enforcement in this area, including in the use of automation, to counter the potential for bias or discrimination, and how to reflect this in transparency reporting and/or enable independent researchers access to relevant data.*

Prior decisions

The Board has previously engaged with Dangerous Individuals and Organizations policy in several cases, including

- *[Video after Nigeria church attack](#), December 2022, (see Meta's response [here](#))*
- *[Mention of the Taliban in news reporting](#), September 2022, (see Meta's response [here](#))*
- *[Colombian Police Cartoon](#), September 2022, (see Meta's response [here](#))*
- *[Shared Al Jazeera post](#), September 2021, (see Meta's response [here](#))*
- *[Öcalan's isolation](#), July 2021, (see Meta's response [here](#))*
- *[Former President Trump's suspension](#), May 2021, (see Meta's response [here](#))*
- *[Punjabi concern over the RSS in India](#), April 2021, (see Meta's response [here](#))*
- *[Nazi quote](#), January 2021, (see Meta's response [here](#))*

Public comments

An important part of the Board's process for developing a policy advisory opinion is gathering additional insights and expertise from individuals and organizations. This input will allow Board members to tap into more knowledge and understand how Meta's policies affect different people in different parts of the world.

If you or your organization feel that you can contribute valuable perspectives to this request for a policy advisory opinion, you can submit your contributions [here](#).

The public comment window for the policy advisory opinion request announced today is open until 15:00 UTC on Monday 10 April.

This timeline is longer than the public comment period for new cases as the policy advisory opinion process does not have the same time constraints as case decisions. Additionally, public comments can be up to six pages in length. For this policy advisory opinion, comments can be submitted in Bengali, as well as any of the languages available on the Board's website, which include Arabic, Farsi, Urdu, Bahasa Indonesia, and Turkish. The full list of languages is available through the link above.

What's next

Now that the Board has accepted this request for a policy advisory opinion, it is collecting necessary information, including through the call for public comments, which has launched today. Following deliberations, the full Board will then vote on the policy advisory opinion. If approved, this will be published on the [Board's website](#).