

# POLICY FORUM

∞ Meta

October 10, 2023

# Reviewing DOI Types, Tiers, and Enforcement

Core Policy Team

# Reviewing DOI Types, Tiers, and Enforcement

## Issue Statement:

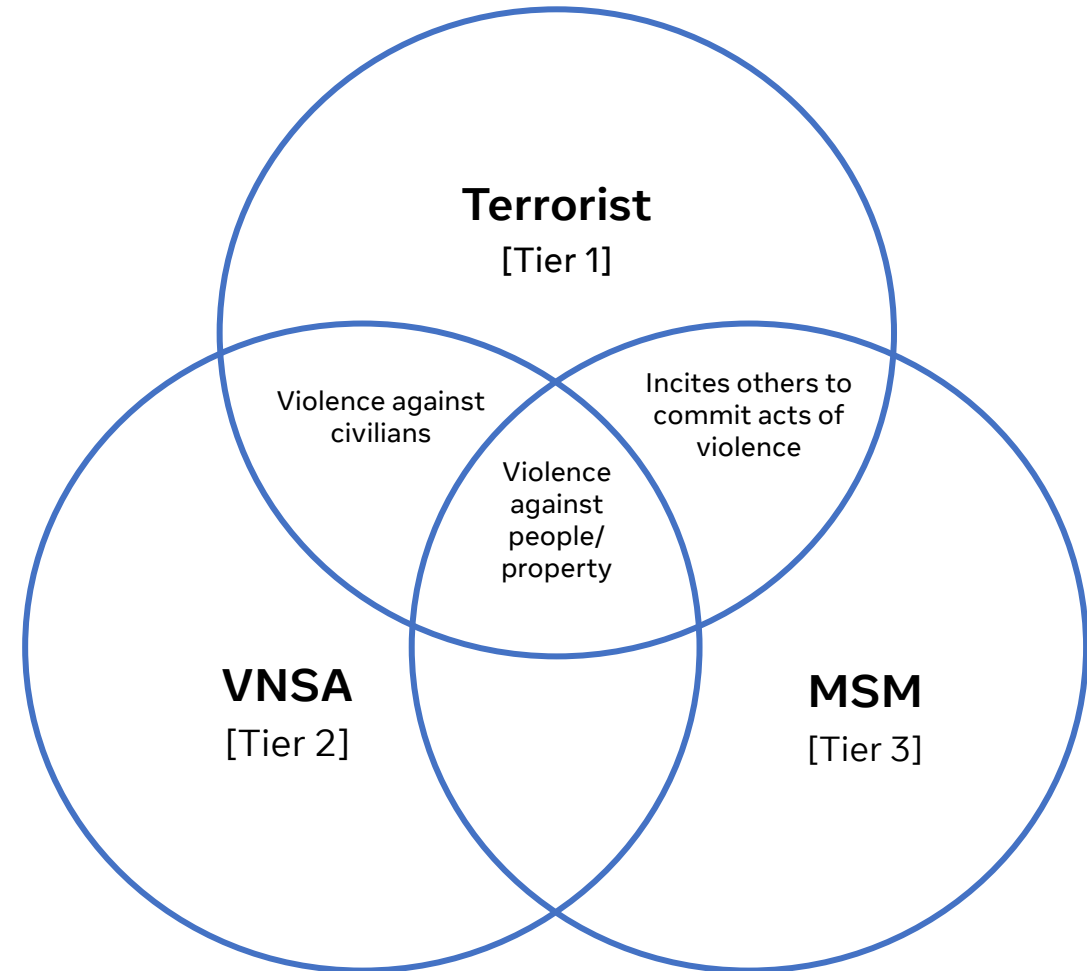
Our Dangerous Organizations and Individuals (DOI) policy categorizes dangerous actors into three enforcement tiers depending on their ties to offline harm and violence. However, policy complexity results in enforcement outcomes that are misaligned with the relative risk posed by the content. We are examining whether a streamlined policy approach would protect benign voice while removing content that glorifies and supports dangerous actors.

| Tiers  | Types  | Enforcement   |
|--------|--|---|
| TIER 1 | Terrorism                                    | Remove References, Glorification, Support, and Representation |
|        | Hate Organizations                           |   |
|        | Criminal Organizations                       |   |
| TIER 2 | Violent Non-State Actors (VNSA)              | Remove Glorification, Support and Representation              |
| TIER 3 | Militarized Social Movements (MSM)           | Remove Representation   |
|        | Violence Inducing Conspiracy Networks (VICN) | Remove Representation (with additional context)               |
|        | Hate-Banned Entities                         | Remove Representation (with additional context)               |

# Problem Analysis

## Types

Overlapping designation signals



# Problem Analysis

## Types

Overlapping designation signals

## Tiering and Enforcement

**Policy complexity leading to inefficient enforcement**

**Lack of user notification and education**

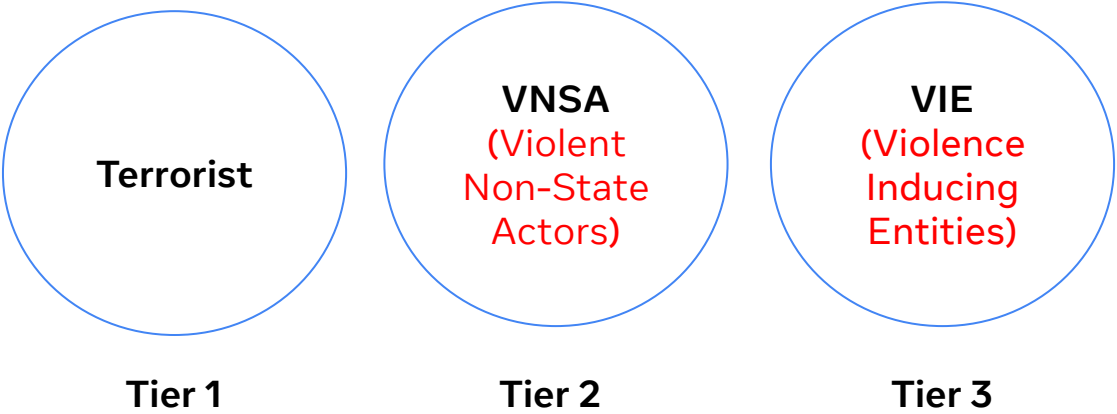
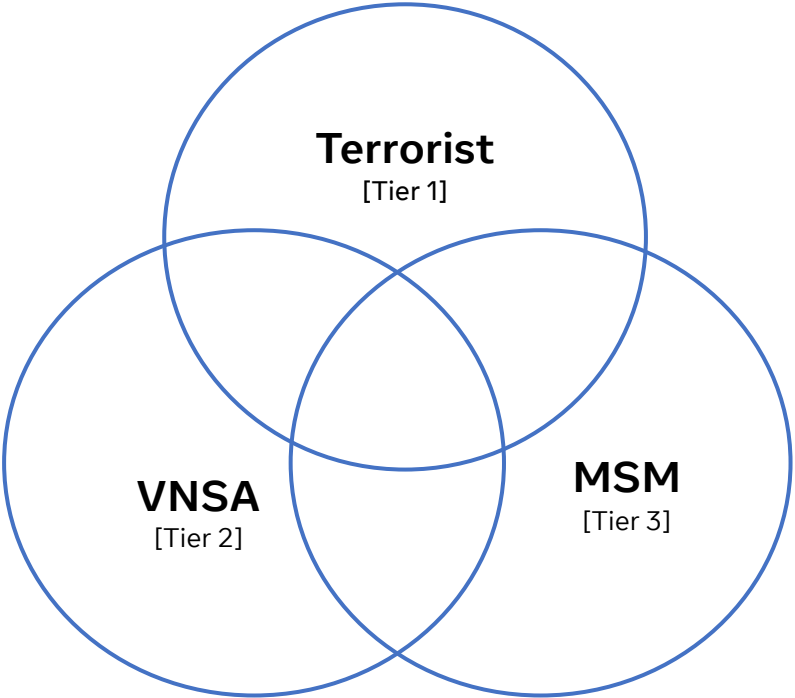
## External Engagement

- Stakeholders favored a **simplified tiering and enforcement system** for effective listing and enforcement but disagreed on the lines across DOI types
- Experts highlighted the **fluid relationship between various tiers**, noting that Tier 1 and 2 entities often graduate up from Tier 3
- Experts recommended paying attention to **coded and implicit signals and region specific behaviors** for better detection, but acknowledged challenges with operationalization.

## Research

- There is little consensus about how to discern across general attributes and motivation(s) of violent entities, and **designations are contestable**.
  - For hate groups, people in 12 countries largely see Meta as an appropriate decider.
- In some contexts, online glorification of violent entities may lead to offline violence, but this pattern **may not generalize** to the range of entities that we classify as Tier 2 and 3 DOI.

# Solutions - Distinct DOI Types & Unified Tier 3



# Options for Tiering and Enforcement

| Option 1: Status Quo |  |   |
|----------------------|--|---|
| Tiers                | Types                                      | Enforcement   |
| Tier 1               | Terrorists,<br>Criminal Orgs,<br>Hate Orgs | Remove References, Glorification, Support, and Representation |
| TIER 2               | VNSA                                       | Remove Glorification, Support and Representation              |
| TIER 3               | MSM  | Remove Representation   |
|                      | VICN                                       | Remove Representation<br><i>(with additional context)</i>     |
|                      | HBE  | Remove Representation<br><i>(with additional context)</i>     |



## Options for Tiering and Enforcement

| Option 2: Aligned Enforcement Across Tier 3 DOI Types |       |  |
|---|-------|--|
| Tiers   | Types | Enforcement                                      |
| TIER 2  | VNSA  | Remove Glorification, Support and Representation |
| TIER 3  | VIE   | Remove Representation                            |

### What's changing between Option 1 and Option 2?

Sub-tiers within Tier 3 are combined into VIE

Remove Representation of all Tier 3 entities

## Recommended Option

| Option 3: Distinguishing between Material and Other Support |           |  |
|---|-----------|--|
| Tier  | Types     | Enforcement  |
| TIER 2  | VNSA, VIE | Remove Glorification, Material Support, Representation |

### What's changing between Option 2 and Option 3?

Combining Tier 2 and Tier 3 into one tier with uniform enforcement

Removing Glorification and Material Support

Allowing Other [Non-Violent] Support

# Options for Tiering and Enforcement

| Option 4: Removing Material and Other Support for new Tier 2 |           |   |
|--|-----------|---|
| Tiers  | Types     | Enforcement                                   |
| TIER 2   | VNSA, VIE | Remove Glorification, Support, Representation |

**What's changing between Option 3 and Option 4?**

Removing all Support

# Options Comparison

| Option 1: Status Quo |       |   |
|----------------------|-------|---|
| Tiers                | Types | Enforcement Action  |
| TIER 2               | VNSA  | Remove Glorification, Support and Representation          |
| TIER 3               | MSM   | Remove Representation                                     |
|                      | VICN  | Remove Representation<br><i>(with additional context)</i> |
|                      | HBE   | Remove Representation<br><i>(with additional context)</i> |

| Option 2: Aligned Enforcement Across Tier 3 DOI Types |            |  |
|---|------------|--|
| Tiers   | Types      | Enforcement Action                               |
| TIER 2  | VNSA       | Remove Glorification, Support and Representation |
| <b>TIER 3</b>   | <b>VIE</b> | <b>Remove Representation</b>                     |

| [Recommendation] Option 3: Distinguishing between Material and Other Support |                  |   |
|--|------------------|---|
| Tiers  | Types            | Enforcement Action  |
| <b>TIER 2</b>  | <b>VNSA, VIE</b> | <b>Remove Glorification, Material Support, Representation</b> |

| Option 4: Removing Material and Other Support for new Tier 2 |                  |  |
|--|------------------|--|
| Tiers  | Types            | Enforcement Action                                   |
| <b>TIER 2</b>  | <b>VNSA, VIE</b> | <b>Remove Glorification, Support, Representation</b> |

∞ Meta