1. Violence and Incitement

I. Violence and Criminal Behavior

Policy Rationale

We aim to prevent potential offline harm that may be related to content on Facebook. While we understand that people commonly express disdain or disagreement by threatening or calling for violence in non-serious ways, we remove language that incites or facilitates serious violence. We remove content, disable accounts and work with law enforcement when we believe there is a genuine risk of physical harm or direct threats to public safety. We also try to consider the language and context in order to distinguish casual statements from content that constitutes a credible threat to public or personal safety. In determining whether a threat is credible, we may also consider additional information like a person's public visibility and the risks to their physical safety.

In some cases, we see aspirational or conditional threats directed at terrorists and other violent actors (e.g. "Terrorists deserve to be killed"), and we deem those non-credible, absent specific evidence to the contrary.

Do not post:

Threats that could lead to death (and other forms of high-severity violence) targeting people or places where threat is defined as any of the following:

- Statements of intent to commit high-severity violence. This includes content where a symbol represents the target and/or includes a visual of an armament or method to represent violence
- Calls for high-severity violence including content where no target is specified but a symbol represents the target and/or includes a visual of an armament or method that represents violence
- Statements advocating for high-severity violence
- Aspirational or conditional statements to commit high-severity violence

Content that asks or offers services for hire to kill others (for example, hitmen, mercenaries, assassins) or advocates for the use of a hitman, mercenary or assassin against a target.

Admissions, statements of intent or advocacy, calls to action, or aspirational or conditional statements to kidnap a target.

Content that depicts abductions or kidnappings if it is clear the content is not being shared by a victim or their family as a plea for help, or shared for informational, condemnation or awareness raising purposes

Threats of high-severity violence using digitally produced or altered imagery to target living people with armaments, methods of violence or dismemberment.

Threats that lead to serious injury (mid-severity violence) towards private individuals, unnamed specified persons, minor public figures, high risk persons, or high risk groups where threat is defined as any of the following:

- Statements of intent to commit violence; or
- Statements advocating violence; or
- Calls for mid-severity violence including content where no target is specified but a symbol represents the target; or
- Aspirational or conditional statements to commit violence; or
- Content about other target(s) apart from private individuals, minor public figures, high risk persons, or high risk groups and any credible:
  - Statements of intent to commit violence; or
  - Calls for action of violence; or
  - Statements advocating for violence; or
  - Aspirational or conditional statements to commit violence

Threats that lead to physical harm (or other forms of lower-severity violence) towards private individuals (self-reporting required) or minor public figures where threat is defined as any of the following:

- Private individuals (name and/or face match are required) or minor public figures that includes:
  - Statements of intent, calls for action, advocating, aspirational or conditional statements to commit low-severity violence

Any content created for the express purpose of outing an individual as a member of a designated and recognisable at-risk group

Instructions on how to make or use weapons if there's evidence of a goal to seriously injure or kill people, through:

- Language explicitly stating that goal, or
- Photos or videos that show or simulate the end result (serious injury or death) as part of the instruction
- Unless the aforementioned content is shared as part of recreational self defense, for military training purposes, commercial video games, or news coverage (posted by Page or with news logo)

Providing instructions on how to make or use explosives, unless there is clear context that the content is for a non-violent purpose (for example part of commercial video games, clear scientific/educational purpose, fireworks, or specifically for fishing)

Any content containing statements of intent, calls for action, conditional or aspirational statements, or advocating for violence due to voting, voter registration or the administration or outcome of an election

Statements of intent or advocacy, calls to action, or aspirational or conditional statements to bring weapons to locations, including but not limited to places of

worship, educational facilities polling places, or locations used to count votes or administer an election (or encouraging others to do the same).

For the following Community Standards, we require additional information and/or context to enforce:

Do not post:

- Content that puts LGBTQI+ people at risk by revealing their sexual identity against their will or without permission
- Content that puts unveiled women at risk by revealing their images without veil against their will or without permission
- Violent threats against law enforcement officers
- Violent threats against people accused of a crime. We remove this content when we have reason to believe that the content is intended to cause physical harm
- Misinformation and unverifiable rumors that contribute to the risk of imminent violence or physical harm
    - Additionally, we have specific rules and guidance regarding content related to COVID-19 and vaccines. To see these specific rules please click here.
- Coded statements where the method of violence or harm is not clearly articulated, but the threat is veiled or implicit. Facebook looks at the below signals to determine whether there is a threat of harm in the content.
    - Shared in a retaliatory context (e.g., expressions of desire to do something harmful to others in response to a grievance or threat that may be real, perceived or anticipated)
    - References to historical or fictional incidents of violence (e.g., content that threatens others by referring to known historical incidents of violence that have been executed throughout history or in fictional settings)
    - Acts as a threatening call to action (e.g., content inviting or encouraging others to carry out harmful acts or to join in carrying out the harmful acts)
    - Indicates knowledge of or shares sensitive information that could expose others to harm (e.g., content that either makes note of or implies awareness of personal information that might make a threat of physical violence more credible. This includes implying knowledge of a person's residential address, their place of employment or education, daily commute routes, or current location)
    - Local context or subject matter expertise confirms that the statement in question could be threatening, and/or could lead to imminent violence or physical harm.
    - The subject of the threat reports the content to us.
- Threats against election officials
- Implicit statements of intent or advocacy, calls to action, or aspirational or conditional statements to bring armaments to locations, including but not limited to places of worship, educational facilities, polling places, or locations used to count votes or administer an election (or encouraging others to do the same). We may also restrict calls to bring armaments to certain locations

where there are temporarily signals of a heightened risk of violence or offline harm. This may be the case, for example, when there's a known protest and counter-protest planned or violence broke out at a protest in the same city within the last 7 days

2. Dangerous Individuals and Organizations[1]

I. Violence and Criminal Behavior

Policy Rationale

In an effort to prevent and disrupt real-world harm, we do not allow organizations or individuals that proclaim a violent mission or are engaged in violence to have a presence on Facebook. We assess these entities based on their behavior both online and offline, most significantly, their ties to violence. Under this policy, we designate individuals, organizations, and networks of people. These designations are divided into three tiers that indicate the level of content enforcement, with Tier 1 resulting in the most extensive enforcement because we believe these entities have the most direct ties to offline harm.

Tier 1 focuses on entities that engage in serious offline harms - including organizing or advocating for violence against civilians, repeatedly dehumanizing or advocating for harm against people based on protected characteristics, or engaging in systematic criminal operations. Tier 1 entities include terrorist, hate, and criminal organizations. We remove praise, substantive support, and representation of Tier 1 entities as well as their leaders, founders, or prominent members.

In addition, we do not allow content that praises, substantively supports, or represents events that Facebook designates as violating violent events - including terrorist attacks, hate events, mass murders or attempted mass murders, multiple murders, or hate crimes. Nor do we allow praise, substantive support, or representation of the perpetrator(s) of such attacks. We also remove content that praises, substantively supports or represents ideologies that promote hate, such as nazism and white supremacy.

Tier 2 focuses on entities that engage in violence against state or military actors but do not generally target civilians -- what we call "Violent Non-State Actors." We remove all substantive support and representation of these entities, their leaders, and their prominent members. We remove any praise of these groups' violent activities.

---

[1] Note from Facebook's records: "This update does not make substantive changes to the Dangerous Organizations and Individuals policy, but rather restructures the formatting, consolidates and formalizes adjustments announced in various statements but not listed in the Community Standards, and updates the policy rationale to better reflect our thinking surrounding dangerous organizations and individuals in order to offer additional user transparency."

Tier 3 focuses on entities that may repeatedly engage in violations of our Hate Speech or Dangerous Organizations policies on-or-off the platform or demonstrate strong intent to engage in offline violence in the near future, but have not necessarily engaged in violence to date or advocated for violence against others based on their protected characteristics. This includes Militarized Social Movements, Violence-Inducing Conspiracy Networks, and individuals and groups banned for promoting hatred. Tier 3 entities may not have a presence or coordinate on our platforms.

We recognize that users may share content that includes references to designated dangerous organizations and individuals to report on, condemn, or neutrally discuss them or their activities. Our policies are designed to allow room for these types of discussions, but we require people to clearly indicate their intent. If the intention is unclear, we may remove content.

We Remove:

We remove Praise, Substantive Support, and Representation of various Dangerous Organizations. These concepts apply to the organizations themselves, their activities, and their members. These concepts do not proscribe peaceful advocacy for particular political outcomes.

Praise, defined as any of the below:

- Speak positively about a designated entity or event;
    - E.g., "The fighters for the Islamic State are really brave!"
- Give a designated entity or event a sense of achievement;
    - E.g., "Timothy McVeigh is a martyr."
- Legitimizing the cause of a designated entity by making claims that their hateful, violent, or criminal conduct is legally, morally, or otherwise justified or acceptable;
    - E.g., "Hitler did nothing wrong."
- Aligning oneself ideologically with a designated entity or event.
    - E.g., "I stand with Brenton Tarrant."

We remove Praise of Tier 1 entities and designated events. We will also remove praise of violence carried out by Tier 2 entities.

Substantive Support, defined as any of the below:

- Any act which improves the financial status of a designated entity - including funnelling money towards, or away from a designated entity;
    - E.g., "Donate to the KKK!"
- Any act which provides material aid to a designated entity or event;
    - E.g., "If you want to send care packages to the Sinaloa Cartel, use this address:"
- Putting out a call to action on behalf of a designated entity or event;
    - E.g., "Contact the Atomwaffen Division - (XXX) XXX-XXXX"
- Recruiting on behalf of a designated entity or event;
    - E.g., "If you want to fight for the Caliphate, DM me"

- Channeling information or resources, including official communications, on behalf of a designated entity or event
  - E.g., Directly quoting a designated entity without caption that condemns, neutrally discusses, or is a part of news reporting.

We remove Substantive Support of Tier 1 and Tier 2 entities and designated events.

Representation, defined as any of the below:

- Stating that you are a member of a designated entity, or are a designated entity;
  - E.g., "I am a grand dragon of the KKK."
- Creating a Page, Profile, Event, Group, or other Facebook entity that is or purports to be owned by a Designated Entity or run on their behalf, or is or purports to be a designated event.
  - E.g., A Page named "American Nazi Party."

We remove Representation of Tier 1 and 2 Designated Organizations, Hate-Banned Entities, and designated events.

Types and Tiers of Dangerous Organizations

Tier 1: Terrorism, organized hate, large-scale criminal activity, mass and multiple murderers, and violating violent events

We do not allow individuals or organizations involved in terrorism, organized hate, or organized crime to have a presence on our platforms. We also don't allow other people to represent these entities on our services. We do not allow leaders or prominent members of these organizations to have a presence on the platform, symbols that represent them to be used on the platform, or content that praises them or their acts. In addition, we remove any coordination of substantive support for these individuals and organizations.
We do not allow content that praises, substantively supports, or represents events that Facebook designates as terrorist attacks, hate events, mass murders or attempted mass murders, serial murders, hate crimes or violating violent events.

We also do not allow Praise, Substantive Support, or Representation of designated hateful ideologies.

Terrorist organizations and individuals, defined as a non-state actor that:

- Engages in, advocates, or lends substantial support to purposive and planned acts of violence,
- Which causes or attempts to cause death, injury or serious harm to civilians, or any other person not taking direct part in the hostilities in a situation of armed conflict, and/or significant damage to property linked to death, serious injury or serious harm to civilians
- With the intent to coerce, intimidate and/or influence a civilian population, government, or international organization
- In order to achieve a political, religious, or ideological aim.

Hate Organizations, defined as an association of three or more people that:

- is organized under a name, sign, or symbol; and
- has an ideology, statements, or physical actions that attack individuals based on characteristics, including race, religious affiliation, national origin, disability, ethnicity, gender, sex, sexual orientation, or serious disease.

Criminal Organizations, defined as an association of three or more people that:

- is united under a name, color(s), hand gesture(s) or recognized indicia; and
- has engaged in or threatens to engage in criminal activity such as homicide, drug trafficking, or kidnapping.

Mass and Multiple Murderers

- We consider an event to be a mass murder or an attempted mass murder if it results in three or more casualties in one incident, defined as deaths or serious injuries. Any Individual who has committed such an attack is considered to be a mass murderer or an attempted mass murderer.
- We consider any individual who has committed two or more murders over multiple incidents or locations a multiple murderer.

Hateful Ideologies

- While our designations of organizations and individuals focus on behavior, we also recognize that there are certain ideologies and beliefs that are inherently tied to violence and attempts to organize people around calls for violence or exclusion of others based on their protected characteristics. In these cases, we designate the ideology itself and remove content that supports this ideology from our platform. These ideologies include:

  - Nazism
  - White Supremacy
  - White Nationalism
  - White Separatism
- We remove explicit Praise, Substantive Support, and Representation of these ideologies, and remove individuals and organizations that ascribe to one or more of these hateful ideologies.

Tier 2: Violent Non-State Actors

Organizations and individuals designated by Facebook as Violent Non-state Actors are not allowed to have a presence on Facebook, or have a presence maintained by others on their behalf. As these groups are actively engaged in violence, substantive support of these entities is similarly not allowed. We will also remove praise of violence carried out by these entities.

Violent Non-State Actors, defined as any non-state actor that:

- engages in purposive and planned acts of violence primarily against a government military or other armed groups; and
- that causes or attempts to
  - cause death to persons taking direct part in hostilities in an armed conflict, and/or
  - deprive communities of access to vital infrastructure and natural resources, and/or bring significant damage to property, linked to death, serious injury or serious harm to civilians

Tier 3: Militarized Social Movements, Violence-Inducing Conspiracy Networks, and Hate Banned Entities

Pages, Groups, Events, and Profiles or other Facebook entities that are - or claim to be - maintained by, or on behalf of, Militarized Social Movements and Violence-Inducing Conspiracy Networks are prohibited. Admins of these pages, groups and events will also be removed.

Click here to read more about how we address movements and organizations tied to violence.

We do not allow Representation of Organizations and individuals designated by Facebook as Hate-Banned Entities.

Militarized Social Movements (MSMs), which include:

- **Militia Groups,** defined as non-state actors that use weapons as a part of their training, communication, or presence; and are structured or operate as unofficial military or security forces and:
  - Coordinate in preparation for violence or civil war; or
  - Distribute information about the tactical use of weapons for combat; or
  - Coordinate militarized tactical coordination in a present or future armed civil conflict or civil war.
- **Groups supporting violent acts amid protests**, defined as non-State actors that repeatedly:
  - Coordinate, promote, admit to or engage in:
  - Acts of street violence against civilians or law enforcement; or
  - Arson, Looting, or other destruction of property; or
  - Threaten to violently disrupt an election process; or
  - Promote bringing weapons to a location when the stated intent is to intimidate people amid a protest.

Violence-Inducing Conspiracy Networks (VICNs), defined as a Non-State Actor that:

- Organizes under a name, sign, mission statement, or symbol; and
- Promote theories that attribute violent or dehumanizing behavior to people or organizations that have been debunked by credible sources; and

- Has inspired multiple incidents of real-world violence by adherents motivated by the desire to draw attention to or redress the supposed harms promoted by these debunked theories.

Hate-Banned Entities, defined as entities that engage in repeated hateful conduct or rhetoric, but do not rise to the level of a Tier 1 entity because they have not engaged in or explicitly advocated for violence, or because they lack sufficient connections to previously designated organizations or figures.

3. Coordinating Harm and Promoting Crime

I. Violence and Criminal Behavior

Policy Rationale

In an effort to prevent and disrupt offline harm and copycat behavior, we prohibit people from facilitating, organizing, promoting or admitting to certain criminal or harmful activities targeted at people, businesses, property or animals. We allow people to debate and advocate for the legality of criminal and harmful activities, as well as draw attention to harmful or criminal activity that they may witness or experience as long as they do not advocate for or coordinate harm.

Do not post content that falls into the following categories:

<u>Harm against people</u>

- Depicting, admitting to or promoting acts of physical harm against people, including acts of domestic violence - except when shared in the context of redemption or defense.
- Swatting - specifically statements of intent, calls to action, representing, supporting, advocating for, depicting, admitting to or speaking positively about it.
- Depicting, promoting, advocating for or encouraging participation in a high risk viral challenge.
- Promoting or advocating for harmful miracle cures for health issues.
- Coordinating, depicting, admitting to or promoting active and deliberate spread of communicable diseases by you or your associates

<u>Harm against animals</u>

Statements of intent, calls to action, representing, supporting or advocating for, or depicting, admitting to or promoting:

- Acts of physical harm against animals committed by you or your associates except in cases of hunting, fishing, religious sacrifice, food preparation or processing, pest or vermin, self-defense or redemption
- Staged animal vs. animal fights, including acts committed by a third party

<u>Harm against property</u>

Statements of intent, calls to action, representing, supporting or advocating for harm against property that depicts, admits to, or promotes the following acts committed by you or your associates:

- Vandalism
- Hacking when the intent is to hijack a domain, corrupt or disrupt cyber systems, seek ransoms, or gain unauthorized access to data systems
- Theft when committed by you or your associates, as well as positive statements about theft when committed by a third party

Voter and/or Census Fraud

- Offers to buy or sell votes with cash or gifts
- Statements that advocate, provide instructions, or show explicit intent to illegally participate in a voting or census process

Voter and/or Census Interference

- Misrepresentation of the dates, locations, and times, and methods for voting or voter registration or census participation
- Misrepresentation of who can vote, qualifications for voting, whether a vote will be counted, and what information and/or materials must be provided in order to vote.
- Misrepresentation of who can participate in the census and what information or materials must be provided in order to participate
- Content claiming that the U.S. Immigration and Customs Enforcement (ICE) is at a voting location
- Explicit claims that people will be infected by COVID (or another communicable disease) if they participate in the voting process.

For the following content, we include a label so that people are aware the content may be sensitive:

- Imagery depicting a high risk viral challenge if shared with a caption that condemns or raise awareness of the associated risks

For the following Community Standards, we require additional information and/or context to enforce:

Do not post:

- Content revealing the identity of someone as a witness, informant, activist, or individuals whose identity or involvement in a legal case has been restricted from public disclosure
- Imagery that is likely to deceive the public as to its origin if:
    - The entity depicted or an authorized representative objects to the imagery, and
    - The imagery has the potential to cause harm to members of the public

- Other misrepresentations related to voting in an official election or census participation may be subject to false news standards, as referenced in <u>section 20</u>
- Misrepresentation of whether a candidate is running or not
- Calls for coordinated interference that would affect an individual's ability to participate in an official census or election
- Content stating that census or voting participation may or will result in law enforcement consequences (for example, arrest, deport or imprisonment)
- Misrepresentation of government involvement in the census, including, where applicable, that an individual's census information will be shared with another (non-census) government agency
- Statements of intent, support, or advocacy to go to an election site, voting location, or vote counting location when the purpose of going to the site is to monitor or watch voters or election officials' activity using militaristic language (e.g. "war," "army," or "soldier") or an expressed goal to intimidate, exert control or display power (e.g. "let's show them who's boss!," "if they're scared, they won't vote!")
- Content coordinating interference with the administration of the COVID-19 vaccine
- Content calling to action, advocating, or promoting that others not get the COVID-19 vaccine

4. Restricted Goods and Services

I. Violence and Criminal Behavior

Policy Rationale

To encourage safety and deter potentially harmful activities, we prohibit attempts by individuals, manufacturers, and retailers to purchase, sell, raffle, gift, transfer or trade certain goods and services on our platform. Brick-and-mortar and online retailers may promote firearms, alcohol, and tobacco items available for sale off of our services; however, we restrict visibility of this content for minors. We allow discussions about the sale of these goods in stores or by online retailers, as well as advocating for changes to regulations of goods and services covered in this policy.

Do not post:

Firearms

Content that

- Attempts to buy, sell, trade, donate, gift or solicit firearms, firearm parts, ammunition, explosives, or lethal enhancements between private individuals, unless posted by a real brick-and-mortar store, legitimate website, brand or government agency (e.g. police department, fire department)
- Provides access to 3D printing or computer aided manufacturing instructions for firearms or firearms parts even when posted by brick and mortar stores, brands and legitimate websites.

Non-medical drugs:

Content that

- Attempts to buy, sell, trade, donate, gift, or solicit non-medical drugs
- Admits to buying or trading non-medical drugs by the poster of the content by themselves or through others
- Admits to personal use without acknowledgment of or reference to recovery, treatment, or other assistance to combat usage
- Speaks positively, encourages, coordinates or provides instructions for use or make of non-medical drugs

Marijuana and pharmaceutical drugs:

- Content that attempts to buy, sell, trade, donate, gift, or solicit marijuana or pharmaceutical drugs

Alcohol / Tobacco:

- Content that attempts to buy, sell, trade, donate or gift alcohol or tobacco products and is not posted by a Page, Group, or Instagram Profile representing a real brick and mortar store, legitimate website or brand

Endangered species:

Content that

- Attempts to buy, sell, trade, donate, gift, or solicit endangered species or their parts
- Admits to poaching, buying or trading of endangered species and their parts committed by the poster of the content either by themselves or through others
- Depicts poaching of endangered species and their parts committed by the poster of the content by themselves or through others
- Speaks positively, coordinates or encourages the poaching of endangered species and their parts
- Provides instructions to use or make products from endangered species and their parts

Live non-endangered animals excluding livestock:

- Content that attempts to buy, sell or trade live non-endangered animals between private individuals.

Human blood:

- Content that attempts to buy, sell, or trade human blood
- Content that solicits human blood unless for a donation or a gift

Weight loss products:

- Content about weight loss that contains a miracle claim and attempts to buy, sell, trade, donate, or gift weight loss products or supplements

Historical artifacts:

- Content that attempts to buy, sell, trade, donate, gift or solicit historical artifacts

For the following content, we restrict visibility to adults twenty-one years of age and older:

Firearms

- Content posted by or promoting a brick-and-mortar store, legitimate website, brand, or government agency which attempts to buy, sell, trade, donate or gift firearms, firearm parts, ammunition, explosives, or lethal enhancements.

For the following content, we restrict visibility to adults eighteen years of age and older:

Alcohol / Tobacco:

- Content posted by or promoting a brick and mortar store, legitimate website or brand, which attempts to buy, sell, trade, donate or gift of alcohol or tobacco products off of Facebook's services

Bladed weapons:

- Content which attempts to buy, sell, trade, donate or gift bladed weapons.

Weight loss products:

Content that

- Attempts to buy, sell, trade, donate or gift weight loss products
- Admits to or depicts using a weight loss product, unless in condemnation
- Speaks positively, coordinates, or encourages the use of weight loss products
- Provides instructions to use or make weight loss products

Potentially dangerous cosmetic procedures:

Content that

- Attempts to buy, sell, trade, donate or gift a potentially dangerous cosmetic procedure
- Admits to or depicts using a potentially dangerous cosmetic procedure, unless in condemnation
- Speaks positively, coordinates, or encourages the use of a potentially dangerous cosmetic procedure
- Provides instructions to use or perform a potentially dangerous cosmetic procedure

5. Fraud and Deception

I. Violence and Criminal Behavior

Policy Rationale

In an effort to prevent and disrupt harmful or fraudulent activity, we remove content aimed at deliberately deceiving people to gain an unfair advantage or deprive another of money, property, or legal right. However, we allow people to raise awareness and educate others as well as condemn these activities using our platform.

Do not post:

Content that engages in, promotes, encourages, facilitates, or admits to the following activities:

>
> Offering, solicitation and trade of:
>
>> Goods/property purchased with stolen financial information
>>
>> Future exam papers or answer sheets where answers are not meant to be widely shared
>>
>> Educational and professional certificates
>>
>> Credentials for paid subscription services
>>
>> Fake user reviews
>>
>> Fake or manipulated documents, e.g. coupons, medical prescriptions
>>
>> Fake or counterfeit currency
>
> Offering or selling of personally identifiable information
> Coordination or facilitation of betting manipulation, e.g. match fixing
> Recruitment of workforce for the purpose of fraudulent activities, e.g. Social Security Number (SSN) scams
> Other forms of deception aimed at depriving people of money or property such as:
>
>> Confidence schemes, e.g. romance/military impersonation scams
>>
>> Ponzi schemes or Pyramid schemes
>>
>> Setting up false businesses or entities
>>
>> Investment schemes with promise of high rates of return

For the following Community Standards, we require additional information and/or context to enforce:

Do not post:

Content that engages in, promotes, encourages, facilitates, or admits to the following activities:

Bribery
Embezzlement
Money laundering (concealment of the origins of criminally obtained money)

6. Suicide and Self-Injury

II. Safety

Policy Rationale

We care deeply about the safety of the people who use our apps. We regularly consult with experts in suicide and self-injury to help inform our policies and enforcement, and work with organizations around the world to provide assistance to people in distress.

While we don't allow people to intentionally or unintentionally celebrate or promote suicide or self-injury, we do allow people to discuss these topics because we want Facebook to be a space where people can share their experiences, raise awareness about these issues, and seek support from one another.

We define self-injury as the intentional and direct injuring of the body, including self-mutilation and eating disorders. We remove any content that encourages suicide or self-injury, including fictional content such as memes or illustrations and any self-injury content which is graphic, regardless of context. We also remove content that identifies and negatively targets victims or survivors of suicide or self-injury seriously, humorously or rhetorically, as well as real time depictions of suicide or self-injury. Content about recovery of suicide or self-harm that is allowed, but may contain imagery that could be upsetting, such as a healed scar, is placed behind a sensitivity screen.

When people post or search for suicide or self-injury- related content, we will direct them to local organizations that can provide support and if someone is at immediate risk of harming themselves, we will contact local emergency services to get them help.

With respect to live content, experts have told us that if someone is saying they intend to attempt suicide on a livestream, we should leave the content up for as long as possible, because the longer someone is talking to a camera, the more opportunity there is for a friend or family member to call emergency services.

However, to minimize the risk of others being negatively impacted by viewing this content, we will stop the livestream at the point at which the threat turns into an attempt. As mentioned above, in any case, we will contact emergency services if we identify someone is at immediate risk of harming themselves.

Do not post:

Content that promotes, encourages, coordinates, or provides instructions for

   Suicide
   Self-injury
   Eating disorders

Content that depicts graphic self-injury imagery

Except in limited situations of newsworthiness, it is against our policies to post content depicting a person who engaged in a suicide attempt or death by suicide

Content that focuses on depiction of ribs, collar bones, thigh gaps, hips, concave stomach, or protruding spine or scapula when shared together with terms associated with eating disorders.

Content that contains instructions for drastic and unhealthy weight loss when shared together with terms associated with eating disorders.

For the following content, we restrict content to adults over the age of 18, and include a sensitivity screen so that people are aware the content may be upsetting:

   Photos or videos depicting a person's death by suicide that are determined to be newsworthy
   Photos or videos depicting a person who engaged in euthanasia/assisted suicide in a medical setting

For the following content, we include a sensitivity screen so that people are aware the content may be upsetting to some:

   Content that depicts older instances of self-harm such as healed cuts or other non-graphic self-injury imagery in a context of recovery
   Content that depicts ribs, collar bones, thigh gaps, hips, concave stomach, or protruding spine or scapula in a recovery context.

We provide resources to people who post written or verbal admissions of engagement in self injury, including:

   Suicide
   Euthanasia/assisted suicide
   Self-harm
   Eating disorders

For the following Community Standards, we require additional information and/or context to enforce:

We may remove suicide notes when we have confirmation of a suicide or suicide attempt. We try to identify suicide notes using several factors, including but not

limited to, family or legal representative requests, media reports, law enforcement reports or other third party sources (e.g., government agencies, NGOs).

7. Child Sexual Exploitation, Abuse and Nudity[2]

II. Safety

Policy Rationale

We do not allow content that sexually exploits or endangers children. When we become aware of apparent child exploitation, we report it to the National Center for Missing and Exploited Children (NCMEC), in compliance with applicable law. We know that sometimes people share nude images of their own children with good intentions; however, we generally remove these images because of the potential for abuse by others and to help avoid the possibility of other people reusing or misappropriating the images.

We also work with external experts, including the Facebook Safety Advisory Board, to discuss and improve our policies and enforcement around online safety issues, especially with regard to children. Learn more about the technology we're using to fight against child exploitation.

Do not post:

Child sexual exploitation

Content that threatens, depicts, praises, supports, provides instructions for, makes statements of intent, admits participation in or shares links of the sexual exploitation of children (real or non-real minors, toddlers or babies), including but not limited to:

- Sexual intercourse
    - Explicit sexual intercourse or oral sex, defined as mouth or genitals entering or in contact with another person's genitals or anus, where at least one person's genitals are nude.
    - Implied sexual intercourse or oral sex, including when contact is imminent or not directly visible.
    - Stimulation of genitals or anus, including when activity is imminent or not directly visible.
    - Presence of by-products of sexual activity.
    - Any of the above involving an animal.
- Children with sexual elements, including but not limited to:
    - Restraints.

---

[2] Note from Facebook's records: "These edits do not reflect policy changes, rather additions to the Community Standards for transparency and to better reflect our enforcement."

- Focus on genitals.
- Presence of aroused adult.
- Presence of sex toys.
- Sexualised costume.
- Stripping.
- Staged environment (for example, on a bed) or professionally shot (quality/focus/angles).
- Open-mouth kissing.
- Content of children in a sexual fetish context.
- Content that supports, promotes, advocates or encourages participation in pedophilia unless it is discussed neutrally in an academic or verified health context.
- Content that identifies or mocks alleged victims of child sexual exploitation by name or image.

Solicitation

Content that solicits

- Imagery of child sexual exploitation
- Nude imagery of children
- Sexualized imagery of children
- Real-world sexual encounters with children

Inappropriate interactions with children

Content that constitutes or facilitates inappropriate interactions with children, such as:

- Arranging or planning real-world sexual encounters with children
- Purposefully exposing children to sexually explicit language or sexual material
- Engaging in implicitly sexual conversations in private messages with children
- Obtaining or requesting sexual material from children in private messages

Exploitative intimate imagery and sextortion

Content that attempts to exploit minors by:

- Coercing money, favors or intimate imagery with threats to expose intimate imagery or information.
- Sharing, threatening or stating an intent to share private sexual conversations or intimate imagery.

Sexualization of children

Content that solicits imagery of child sexual exploitation, or nude or sexualized images or videos of children

- Content (including photos, videos, real-world art, digital content, and verbal depictions) that shows children in a sexualized context.
- Groups, Pages and profiles targeted to sexualizing children.

Child nudity

Content that depicts child nudity where nudity is defined as:

- Close-ups of children's genitalia
- Real nude toddlers, showing:
    - Visible genitalia, even when covered or obscured by transparent clothing.
    - Visible anus and/or fully nude close-up of buttocks.
- Real nude minors, showing:
    - Visible genitalia (including genitalia obscured only by pubic hair or transparent clothing)
    - Visible anus and/or fully nude close-up of buttocks.
    - Uncovered female nipples.
    - No clothes from neck to knee - even if no genitalia or female nipples are showing.
- Digitally-created depictions of nude minors, toddlers or babies unless the image is for health or educational purposes.

Non-sexual child abuse

Imagery that depicts non-sexual child abuse regardless of sharing intent

Content that praises, supports, promotes, advocates for, provides instructions for or encourages participation in non-sexual child abuse.

For the following content, we include a warning screen so that people are aware the content may be disturbing and limit the ability to view the content to adults, ages eighteen and older:

- Videos or photos that depict police officers or military personnel committing non-sexual child abuse.
- Imagery of non-sexual child abuse, when law enforcement, child protection agencies, or trusted safety partners request that we leave the content on the platform for the express purpose of bringing a child back to safety.

For the following content, we include a sensitivity screen so that people are aware the content may be upsetting to some:

- Videos or photos of violent immersion of a child in water in the context of religious rituals.

For the following Community Standards, we require additional information and/or context to enforce:

For the following content, we include a warning label so that people are aware that the content may be sensitive:

- Imagery posted by a news agency that depicts child nudity in the context of famine, genocide, war crimes, or crimes against humanity, unless accompanied by a violating caption or shared in a violating context, in which case the content is removed.

8. Sexual Exploitation of Adults

II. Safety

Policy Rationale

We recognize the importance of Facebook as a place to discuss and draw attention to sexual violence and exploitation. In an effort to create space for this conversation and promotea safe environment, we allow victims to share their experiences, butremove content that depicts, threatens or promotes sexual violence, sexual assault, or sexual exploitation, . We alsoremove content that displays, advocates for, or coordinates sexual acts with non-consenting parties to avoid facilitating non-consensual sexual acts

To protect victims and survivors, we also remove images that depict incidents of sexual violence and intimate images shared without the consent of the person(s)pictured. As noted in the introduction, we also work with external safety experts to discuss and improve our policies and enforcement around online safety issues, and we may remove content when they provide information that content is linked to harmful activity. We've written about the technology we use to protect against intimate images and the research that has informed our work. We've also put together a guide to reporting and removing intimate images shared without your consent.

Do not post:

In instances where content consists of any form of non-consensual sexual touching, crushing, necrophilia or bestiality,or forced stripping, including:

    Depictions (including real photos/videos), or
    Advocacy (including aspirational and conditional statements), or
    Statements of intent, or
    Calls for action, or
    Threatening, soliciting, or stating an intent to share imagery, or

Admitting participation, or
Mocking victims of any of the above

Content that attempts to exploit people by any of the following:

Sextortion: Coercing money, favors, or intimate imagery from people with threats to expose their intimate imagery or intimate information

Sharing, threatening, stating an intent to share, offering or asking for imagery known as "Revenge Porn" or "Non-Consensual Intimate Images" that fulfills all three of the following conditions:

Image is non-commercial or produced in a private setting

Person in the image is (near) nude, engaged in sexual activity, or in a sexual pose

Lack of consent to share the image is indicated by

Vengeful context (for example, caption, comments, or page title)

Independent sources (for example, law enforcement record) including entertainment media (for example, leak of images confirmed by media)

A visible match between the person depicted in the image and the person who has reported the content to us

The person who reported the content to us shares the same name as the person depicted in the image

Secretly taken non-commercial imagery of a real person's commonly sexualized body parts (breasts, groin, buttocks, or thighs) or of a real person engaged in sexual activity. This imagery is commonly known as "creepshots" or "upskirts" and includes photos or videos that mocks, sexualizes or exposes the person depicted in the imagery.

Threatening or stating an intent to share private sexual conversations that meets the following criteria:

Lack of consent is indicated by:

Vengeful context and/or threatening context, or

A visible match between the person depicted in the image and the person who has reported the content to us

The person who reported the content to us shares the same name as the person depicted in the image

For the following content, we include a warning screen so that people are aware the content may be disturbing:

Narratives and statements that contain a depiction of non-consensual sexual touching (written or verbal) that includes details beyond mere naming or mentioning the act if:

Shared by the victim, or
Shared by a third party (other than the victim) in support of victim or condemnation of act or for general awareness to be determined by context/caption

Content mocking the concept of non-consensual sexual touching

For the following Community Standards, we require additional information and/or context to enforce:

We may restrict visibility to people over the age of 18 and include a warning label on certain fictional videos (e.g., digital and hand-drawn imagery, film and TV show scenes, comic books, music videos) that depict non-consensual sexual touching.

In addition to our at-scale policy of removing content that threatens or advocates rape or other non-consensual sexual touching, we may also disable the posting account.

9. Bullying and Harassment[3]

II. Safety

---

[3] Note from Facebook's records: "In response to an Oversight Board recommendation we committed to reformat the Bullying and Harassment Community Standard to clearly define the approach to different target user categories. Edits to this section have been made to reformat the section accordingly and bring alignment with internal documentation for transparency. We're also updating our policy on mass harassment which was based on global engagement with a broad range of stakeholders directly impacted by the policy, including feminist and women's rights activists, representatives of the LGBT community, minority groups, journalists, human rights activists, and public figures. We also connected with experts who study online harassment and state-sponsored influence operations, as well as advocates for freedom of expression. In our consultations, we explored the way stakeholders experienced brigading and mass harassment on our platforms, including the actors and targets involved, the tactics employed, and the coordinated nature of such attacks. We also discussed at length the numerous online and offline harms that may stem from such behaviors as well as the challenges in differentiating between legitimate forms of mass activity and abusive forms of brigading. Overall, stakeholders acknowledged that mass activity used for activism versus for harmful brigading can often exhibit similar online behaviors (i.e. mass reporting, comment flooding, or hashtag bombing). It was therefore recommended that we focus on context-specific factors to differentiate between legitimate and illegitimate use-cases, such as the nature of the content posted, the impact on the victim, and the behavior's potential for offline harm. The input received on such issues was used to craft the first iteration of our brigading and mass harassment policy."

Policy Rationale

Bullying and harassment happen in many places and come in many different forms from making threats and releasing personally identifiable information to sending threatening messages and making unwanted malicious contact. We do not tolerate this kind of behavior because it prevents people from feeling safe and respected on Facebook.

We distinguish between public figures and private individuals because we want to allow discussion, which often includes critical commentary of people who are featured in the news or who have a large public audience. For public figures, we remove attacks that are severe as well as certain attacks where the public figure is directly tagged in the post or comment.

For private individuals, our protection goes further: We remove content that's meant to degrade or shame, including, for example, claims about someone's sexual personal activity. We recognize that bullying and harassment can have more of an emotional impact on minors, which is why our policies provide heightened protection for users between the ages of 13 and 18.

Context and intent matter, and we allow people to post and share if it is clear that something was shared in order to condemn or draw attention to bullying and harassment. In certain instances, we require self-reporting because it helps us understand that the person targeted feels bullied or harassed. In addition to reporting such behavior and content, we encourage people to use <u>tools available on Facebook</u> to help protect against it.

We also have a <u>Bullying Prevention Hub</u>, which is a resource for teens, parents, and educators seeking support for issues related to bullying and other conflicts. It offers step-by-step guidance, including information on how to start important conversations about bullying. Learn more about what we are doing to protect people from bullying and harassment <u>here</u>.

Note: This policy does not apply to individuals who are part of designated organizations under the <u>Dangerous Organizations and Individuals policy</u> or individuals who died prior to 1900.

Do not:

Tier 1: Target anyone maliciously by:

- Repeatedly contacting someone in a manner that is:
    - Unwanted, or
    - Sexually harassing, or
    - Directed at a large number of individuals with no prior solicitation.

- Attacking someone based on their status as a victim of sexual assault, sexual exploitation, sexual harassment, or domestic abuse.
- Calling for self-injury or suicide of a specific person, or group of people.

- Attacking someone through derogatory terms related to sexual activity (e.g. whore, slut).
- Posting content about a violent tragedy, or victims of violent tragedies that include claims that a violent tragedy did not occur.
- Posting content about victims or survivors of violent tragedies or terrorist attacks by name or by image, with claims that they are:
    - Acting/pretending to be a victim of an event.
    - Otherwise paid or employed to mislead people about their role in the event.

- Threatening to release an individual's private phone number, residential address or email address.
- Making statements of intent to engage in a sexual activity or advocating for them to engage in a sexual activity.
- Creating Pages or Groups that are dedicated to attacking individual(s) by:
    - Calling for death, or to contract or develop a medical condition.
    - Making statements of intent of advocating to engage in sexual activity.
    - Calling for, or making statements of intent to engage in, bullying and/or harassment.
    - Making claims that the individual has or may have a sexually transmitted disease.

- Sending messages that contain the following attacks when aimed at an individual or group of individuals in the thread:
    - Attacks referenced in Tier 1, 2 and 4 of this policy.
    - Targeted cursing.
    - Calls for death, serious disease, disability, epidemic disease or physical harm.

Tier 2: Target private individuals, involuntary public figures or public figures who are minors with:

- Calls for death, or to contract or develop a medical condition.
- Female-gendered cursing terms when used in a derogatory way.
- Claims about sexual activity or sexually transmitted diseases except in the context of criminal allegations against adults about non-consensual sexual touching. .
- Pages or Groups created to attack through:
    - Targeted cursing.
    - Negative physical descriptions.
    - Claims about religious identity or blasphemy.
    - Expressions of contempt or disgust.
    - Female-gendered cursing terms when used in a derogatory way.

Tier 3: Target public figures by purposefully exposing them to:

- For adults:
    - Calls for death, or to contract or develop a medical condition.
    - Claims about sexually transmitted disease

- Female-gendered cursing terms when used in a derogatory way.
- Content that praises, celebrates or mocks their death or medical condition.

- For minors:
    - Comparisons to animals or insects that are culturally perceived as intellectually or physically inferior or to an inanimate object ("cow," "monkey" "potato").
    - Content manipulated to highlight, circle or otherwise negatively draw attention to specific physical characteristics (nose, ear and so on).

Tier 4: Target private individuals or involuntary public figures with:

- Comparisons to animals or insects that are culturally perceived as intellectually or physically inferior or to an inanimate object ("cow," "monkey" "potato").
- Content manipulated to highlight, circle or otherwise negatively draw attention to specific physical characteristics (nose, ear and so on).
- Attacks through negative physical descriptions.
- Content that ranks individuals on physical appearance or personality.
- Content sexualizing another adult.
- Content that further degrades or expresses disgust toward individuals who are depicted menstruating, urinating, vomiting, or defecating.
- Content that further degrades individuals who are depicted being physically bullied except in self-defense and fight-sport contexts..
- Content that praises, celebrates, or mocks their death or serious physical injury.
- In addition to the above, attacks through Pages or Groups:
    - Negative character or ability claims.
    - First-person voice bullying only if the object targets more than one private individual.

Tier 5: Target private adults (who must self-report) or any private minors or involuntary minor public figures with:

- Targeted cursing.
- Claims about romantic involvement, sexual orientation or gender identity.
- Coordination, advocacy or promotion of exclusion.
- Negative character or ability claims, except in the context of criminal allegations and business reviews against adults. We allow criminal allegations so that people can draw attention to personal experiences or offline events. In cases in which criminal allegations pose off-line harm to the named individual, however, we may remove them.
- Expressions of contempt or disgust, except in the context of criminal allegations against adults.

Tier 6: Target private individuals who are minors with:

- Allegations about criminal or illegal behavior.
- Videos of physical bullying shared in a non-condemning context.

Tier 7: Target private individuals (who must self-report) with:

- First-person voice bullying.
- Unwanted manipulated imagery.
- Comparison to other public, fictional or private individuals on the basis of physical appearance
- Claims about religious identity or blasphemy.
- Comparisons to animals or insects that are not culturally perceived as intellectually or physically inferior ("tiger," "lion").
- Neutral or positive physical descriptions.
- Non-negative character or ability claims.
- Any bullying or harassment violation, when shared in an endearing context.
- Attacks through derogatory terms related to a lack of sexual activity.

We add a cover to this content so people can choose whether to see it:

Videos of physical bullying against minors shared in a condemning context

For the following Community Standards, we require additional information and/or context to enforce:

Do not:

- Post content that targets private individuals through unwanted Pages, Groups and Events. We remove this content when it is reported by the victim or an authorized representative of the victim.
- Create accounts to contact someone who has blocked you.
- Post attacks that use derogatory terms related to female gendered cursing. We remove this content when the victim or an authorized representative of the victim informs us of the content, even if the victim has not reported it directly.
- Post content that would otherwise require the victim to report the content or an indicator that the poster is directly targeting the victim (for example the victim is tagged in the post or comment). We will remove this content if we have confirmation from the victim or an authorized representative of the victim that the content is unwanted.
- Post content praising, celebrating or mocking anyone's death. We also remove content targeting a deceased individual that we would normally require the victim to report.
- Post content calling for or stating an intent to engage in behavior that would qualify as bullying and harassment under our policies. We will remove this content when we have confirmation from the victim or an authorized representative of the victim that the content is unwanted.
- Engage in mass harassment against individuals that targets them based on their decision to take or not take the COVID-19 vaccine with:
    - Statements of mental or moral inferiority based on their decision, or
    - Statements that advocate for or allege a negative outcome as a result of their decision, except for widely proven and/or accepted COVID-19 symptoms or vaccine side effects.
- Remove directed mass harassment, when:
    - Targeting, via any surface, 'individuals at heightened risk of offline harm', defined as:

- Human rights defenders
- Minors
- Victims of violent events/tragedies
- Opposition figures in at-risk countries during election periods
- Government dissidents who have been targeted based on their dissident status
- Ethnic and religious minorities in conflict zones
- Member of a designated and recognizable at-risk group
  - Targeting any individual via personal surfaces, such as inbox or profiles, with:
    - Content that violates the bullying and harassment policies for private individuals or,
    - Objectionable content that is based on a protected characteristic
- Disable accounts engaged in mass harassment as part of either
  - State or state-affiliated networks targeting any individual via any surface.
  - Adversarial networks targeting any individual via any surface with:
    - Content that violates the bullying and harassment policies for private individuals or,
    - Content that targets them based on a protected characteristic, or,
    - Content or behavior otherwise deemed to be objectionable in local context

10. Human Exploitation

II. Safety

Policy Rationale

In an effort to disrupt and prevent harm, we remove content that facilitates or coordinates the exploitation of humans, including human trafficking. We define human trafficking as the business of depriving someone of liberty for profit. It is the exploitation of humans in order to force them to engagein commercial sex, labor, or other activities against their will. It relies on deception, force and coercion, and degrades humans by depriving them of their freedom while economically or materially benefiting others.

Human trafficking is multi-faceted and global; it can affect anyone regardless of age, socioeconomic background, ethnicity, gender, or location. It takes many forms, and any given trafficking situation can involve various stages of development. By the coercive nature of this abuse, victims cannot consent.

While we need to be careful not to conflate human trafficking and smuggling, the two can be related and exhibit overlap. The United Nations defines human smuggling as the procurement or facilitation of illegal entry into a state across international

borders. Without necessity for coercion or force, it may still result in the exploitation of vulnerable individuals who are trying to leave their country of origin, often in pursuit of a better life. Human smuggling is a crime against a state, relying on movement, and human trafficking is a crime against a person, relying on exploitation.

Do not post:

Content or exhibit behavior that may lead to human exploitation, which includes any of the following:

>Sex Trafficking, covering minors and adults
>The sale of children for illegal adoption
>Orphanage Trafficking and Orphanage Volun-tourism
>Forced Marriages
>Labor Exploitation (incl. bonded labor)
>Domestic Servitude
>Non-regenerative Organ Trafficking
>Forced Criminal Activity (e.g. forced begging, forced drug trafficking)
>Recruitment of Child Soldiers

Content geared towards the:

>Recruitment of potential victims through force, fraud, coercion, enticement, deception, blackmail or other non-consensual acts.
>Facilitation of human exploitation by coordinating, transporting, transferring, harboring or brokering of victims prior or during the exploitation.
>Exploitation of humans by promoting, depicting, or advocating for it.

Content that offers or assists in smuggling of humans

11. Privacy Violations and Image Privacy Rights

II. Safety

Policy Rationale

Privacy and the protection of personal information are fundamentally important values for Facebook. We work hard to safeguard your personal identity and information and we do not allow people to post personal or confidential information about yourself or of others.

We remove content that shares, offers or solicits personally identifiable information or other private information that could lead to physical or financial harm, including financial, residential, and medical information, as well as private information obtained from illegal sources.

We also provide people ways to report imagery that they believe to be in violation of their privacy rights.

Do not post:

Content that shares or solicits any of the following private information, either on Facebook or through external links:

Personally identifiable information about yourself or others

Personal identity: identifying individuals through government-issued numbers

National identification number (for example Social Security Numbers (SSN), Passport Number, National Insurance/Health Service Number, Personal Public Service Number (PPS), Individual Taxpayer Identification Numbers (ITINs))

Government IDs of law enforcement, military, or security personnel

Personal information: directly identifying an individual, by indicating the ID number or registration information and the individual's name

Records or official documentation of civil registry information (marriage, birth, death, or name change, gender recognition, and so on)

Immigration and work status documents (for example, green cards, work permits, visas, or immigration papers)

Driver's licenses or license plates

Credit Privacy Number (CPN)

Digital identity: authenticating access to an online identity

Email addresses with passwords

Digital identities with passwords

Passwords, pins, or codes to access private information

Personal contact information of others such as phone numbers, addresses and email addresses

Other Private Information

Financial information

Personal financial information about yourself or others, including:

Non-public financial records or statements

Bank account numbers with security or pin codes

Digital payment method information with log-in details, security or pin codes

Credit or debit card information with validity dates or security pins or codes

Financial information about businesses or organizations, including:

Financial records or statements except when the financial records of the business are publicly available (for example listed on stock exchanges, regulatory agencies, and so on)

Bank account numbers accompanied by security or pin codes

Digital payment method information accompanied by log-in details, security or pin codes

Residential information

Imagery that display the external view of private residences if all of the following conditions apply:

The residence is a single-family home, or the resident's unit number is identified in the image/caption

The city/neighborhood or GPS pins (for example, a pin from Google Maps) are identified

The content identifies the resident(s)

That same resident objects to the exposure of their private residence or there is context of organizing protests against the resident (this does not include embassies that also serve as residences)

Content that exposes information about safe houses by sharing any of the below, unless the safe house is actively promoting information about their facility

Actual address (note: 'Post Box only' is allowed)

Images of the safe house

Identifiable city/neighborhood of the safe house

Information exposing the identity of the safe house residents

Medical information

Records or official documentation displaying medical, psychological, biometric, or genetic hereditary of others

Information obtained from hacked sources

Except in limited cases of newsworthiness, content claimed or confirmed to come from a hacked source, regardless of whether the affected person is a public figure or a private individual

Status or occupation related information

Content that exposes the undercover status of law enforcement, military, or security personnel if:

> The content contains the agent's full name or other explicit identification and explicitly mentions their undercover status

> The imagery identifies the faces of the law enforcement personnel and explicitly mentions their undercover status

The following content also may be removed

A reported photo or video of people where the person depicted in the image is:

> A minor under thirteen years old, and the content was reported by the minor or a parent or legal guardian,

> A minor between thirteen and eighteen years old, and the content was reported by the minor,

> An adult, where the content was reported by the adult from outside the United States and applicable law may provide rights to removal

> Any person who is incapacitated and unable to report the content on their own

For the following Community Standards, we require additional information and/or context to enforce:

We remove:

Content that puts a defector at risk by outing the individual with personally identifiable information when the content is reported by credible government channels

Depictions of someone in a medical or health facility if reported by the person pictured or an authorized representative

12. Hate Speech[4]

---

[4] Note from Facebook's records: "This update was based on stakeholder engagement with a broad range of academics and civil society organizations affected by the policy, including experts in dangerous speech and atrocity prevention, human rights practitioners, social psychologists who study issues of personal identity, advocates for freedom of expression, and groups representing religious and non-religious world views. In our consultations, we explored the harms that may stem from attacks on concepts, ideas, and institutions; we also listened to stakeholders describe the legitimate role of such speech in advancing public discourse and preserving dissent. Speech that was perceived as hurtful by some stakeholder groups was regarded as serving a valuable civic purpose by others. Two

III. Objectionable Content

Policy Rationale

We believe that people use their voice and connect more freely when they don't feel attacked on the basis of who they are. That is why we don't allow hate speech on Facebook. It creates an environment of intimidation and exclusion, and in some cases may promote offline violence.

We define hate speech as a direct attack against people — rather than concepts or institutions— on the basis of what we call protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease. We define attacks as violent or dehumanizing speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation. We also prohibit the use of harmful stereotypes, which we define as dehumanizing comparisons that have historically been used to attack, intimidate, or exclude specific groups, and that are often linked with offline violence. We consider age a protected characteristic when referenced along with another protected characteristic. We also protect refugees, migrants, immigrants and asylum seekers from the most severe attacks, though we do allow commentary and criticism of immigration policies. Similarly, we provide some protections for characteristics like occupation, when they're referenced along with a protected characteristic.

We recognize that people sometimes share content that includes someone else's hate speech to condemn it or raise awareness. In other cases, speech that might otherwise violate our standards can be used self-referentially or in an empowering way. Our policies are designed to allow room for these types of speech, but we require people to clearly indicate their intent. If the intention is unclear, we may remove content.

Learn more about our approach to hate speech.

Do not post:

Tier 1

Content targeting a person or group of people (including all subsets except those described as having carried out violent crimes or sexual offenses) on the basis of their aforementioned protected characteristic(s) or immigration status with:

---

themes stood out: The first was the importance of ensuring that attacks on concepts, ideas, and institutions do not contribute to real world harm. This concern found expression in support for international guidelines on incitement to hatred (e.g., the Rabat Plan of Action). The second was the importance of factors such as history, local power dynamics, the status of the speaker, and ongoing civic tension -- which may tip the scales in favor of removing attacks in some circumstances but not others. We used the input we received on these points to craft a revision to our hate speech policy."

- Violent speech or support in written or visual form
- Dehumanizing speech or imagery in the form of comparisons, generalizations, or unqualified behavioral statements (in written or visual form) to or about:
    - Insects.
    - Animals that are culturally perceived as intellectually or physically inferior.
    - Filth, bacteria, disease and feces.
    - Sexual predator.
    - Subhumanity.
    - Violent and sexual criminals
    - Other criminals (including but not limited to "thieves," "bank robbers," or saying "All [protected characteristic or quasi-protected characteristic] are 'criminals'").
    - Statements denying existence.
- Mocking the concept, events or victims of hate crimes even if no real person is depicted in an image.
- Designated dehumanizing comparisons, generalizations, or behavioral statements (in written or visual form) that include:
    - Black people and apes or ape-like creatures.
    - Black people and farm equipment.
    - Caricatures of Black people in the form of blackface.
    - Jewish people and rats.
    - Jewish people running the world or controlling major institutions such as media networks, the economy or the government.
    - Denying or distorting information about the Holocaust.
    - Muslim people and pigs.
    - Muslim person and sexual relations with goats or pigs.
    - Mexican people and worm-like creatures.
    - Women as household objects or referring to women as property or "objects".
    - Transgender or non-binary people referred to as "it".
    - Dalits, scheduled caste or 'lower caste' people as menial laborers.

Tier 2

Content targeting a person or group of people on the basis of their protected characteristic(s) with:

- Generalizations that state inferiority (in written or visual form) in the following ways:
    - Physical deficiencies are defined as those about:
        - Hygiene, including but not limited to: filthy, dirty, smelly.
        - Physical appearance, including but not limited to: ugly, hideous.
    - Mental deficiencies are defined as those about:
        - Intellectual capacity, including but not limited to: dumb, stupid, idiots.
        - Education, including but not limited to: illiterate, uneducated.

- Mental health, including but not limited to: mentally ill, retarded, crazy, insane.
  - Moral deficiencies are defined as those about:
    - Character traits culturally perceived as negative, including but not limited to: coward, liar, arrogant, ignorant.
    - Derogatory terms related to sexual activity, including but not limited to: whore, slut, perverts.

- Other statements of inferiority, which we define as:
  - Expressions about being less than adequate, including but not limited to: worthless, useless.
  - Expressions about being better/worse than another protected characteristic, including but not limited to: "I believe that males are superior to females."
  - Expressions about deviating from the norm, including but not limited to: freaks, abnormal.

- Expressions of contempt (in written or visual form), which we define as:
  - Self-admission to intolerance on the basis of a protected characteristics, including but not limited to: homophobic, islamophobic, racist.
  - Expressions that a protected characteristic shouldn't exist.
  - Expressions of hate, including but not limited to: despise, hate.

- Expressions of dismissal, including but not limited to: don´t respect, don't like, don´t care for

- Expressions of disgust (in written or visual form), which we define as:
  - Expressions that suggest the target causes sickness, including but not limited to: vomit, throw up.
  - Expressions of repulsion or distaste, including but not limited to: vile, disgusting, yuck.

- Cursing, defined as:
  - Referring to the target as genitalia or anus, including but not limited to: cunt, dick, asshole.
  - Profane terms or phrases with the intent to insult, including but not limited to: fuck, bitch, motherfucker.
  - Terms or phrases calling for engagement in sexual activity, or contact with genitalia, anus, feces or urine, including but not limited to: suck my dick, kiss my ass, eat shit.

Tier 3

Content targeting a person or group of people on the basis of their protected characteristic(s) with any of the following:

- Segregation in the form of calls for action, statements of intent, aspirational or conditional statements, or statements advocating or supporting segregation.

- Exclusion in the form of calls for action, statements of intent, aspirational or conditional statements, or statements advocating or supporting, defined as
  - Explicit exclusion, which means things like expelling certain groups or saying they are not allowed.
  - Political exclusion, which means denying the right to right to political participation.
  - Economic exclusion, which means denying access to economic entitlements and limiting participation in the labour market.
  - Social exclusion, which means things like denying access to spaces (physical and online)and social services.

Content that describes or negatively targets people with slurs, where slurs are defined as words that are inherently offensive and used as insulting labels for the above characteristics.

For the following Community Standards, we require additional information and/or context to enforce:

Do not post:

- Content explicitly providing or offering to provide products or services that aim to change people's sexual orientation or gender identity.
- Content attacking concepts, institutions, ideas, practices, or beliefs associated with protected characteristics, which are likely to contribute to imminent physical harm, intimidation or discrimination against the people associated with that protected characteristic. Facebook looks at a range of signs to determine whether there is a threat of harm in the content. These include but are not limited to: content that could incite imminent violence or intimidation; whether there is a period of heightened tension such as an election or ongoing conflict; and whether there is a recent history of violence against the targeted protected group. In some cases, we may also consider whether the speaker is a public figure or occupies a position of authority.

13. Violent and Graphic Content

III. Objectionable Content

Policy Rationale

We remove content that glorifies violence or celebrates the suffering or humiliation of others because it may create an environment that discourages participation. We allow graphic content (with some limitations) to help people raise awareness about issues. We know that people value the ability to discuss important issues like human rights abuses or acts of terrorism. We also know that people have different sensitivities with regard to graphic and violent content. For that reason, we add a warning label to especially graphic or violent content so that it is not available to people under the age of eighteen and so that people are aware of the graphic or violent nature before they click to see it.

Do not post:

Imagery of People

Videos of people or dead bodies in non-medical settings if they depict

>   Dismemberment
>   Visible internal organs; partially decomposed bodies
>   Charred or burning people unless in the context of cremation or self-immolation when that action is a form of political speech or newsworthy
>   Victims of cannibalism
>   Throat-slitting

Live streams of capital punishment of a person

For the following content, we include a warning screen so that people are aware the content may be disturbing. We also limit the ability to view the content to adults, ages eighteen and older:

Imagery of People

Videos of people or dead bodies in a medical setting if they depict

>   Dismemberment
>   Visible internal organs; partially decomposed bodies
>   Charred or burning people, including cremation or self-immolation when that action is a form of political speech or newsworthy
>   Victims of cannibalism
>   Throat-slitting

Photos of wounded or dead people if they show

>   Dismemberment
>   Visible internal organs; partially decomposed bodies
>   Charred or burning people
>   Victims of cannibalism
>   Throat-slitting

Imagery that shows the violent death of a person or people by accident or murder

Imagery that shows capital punishment of a person

Imagery that shows acts of torture committed against a person or people

Imagery of non-medical foreign objects (such as metal objects, knives, nails) involuntarily inserted or stuck into people causing grievous injury

Imagery of Animals

The following content involving animals:

> Videos depicting humans killing animals if there is no explicit manufacturing, hunting, or food consumption, processing or preparation context
> Imagery of animal to animal fights, when there are visible innards or dismemberment of non-regenerating body, unless in the wild
> Imagery of humans committing acts of torture or abuse against live animals
> Imagery of animals showing wounds or cuts that render visible innards or dismemberment, if there is no explicit manufacturing, hunting, taxidermy, medical treatment, rescue or food consumption, preparation or processing context or the animal is already skinned or with its outer layer fully removed.

For the following content, we include a label so that people are aware the content may be sensitive:

Imagery of non-medical foreign objects voluntarily inserted into people through skin in religious or cultural context

Imagery of visible innards in a birthing context

Imagery of fetuses and newborn babies that show:

> Dismemberment
> Visible innards
> An abortion or abandonment context

Imagery of newborn babies in an abandonment context

Imagery of animals in a ritual slaughter context showing dismemberment, or visible innards, or charring or burning

For the following Community Standards, we require additional information and/or context to enforce:

We remove:

> Videos and photos that show the violent death of someone when a family member requests its removal

14. Adult Nudity and Sexual Activity

III. Objectionable Content

Policy Rationale

We restrict the display of nudity or sexual activity because some people in our community may be sensitive to this type of content. Additionally, we default to removing sexual imagery to prevent the sharing of non-consensual or underage content. Restrictions on the display of sexual activity also apply to digitally created content unless it is posted for educational, humorous, or satirical purposes.

Our nudity policies have become more nuanced over time. We understand that nudity can be shared for a variety of reasons, including as a form of protest, to raise awareness about a cause, or for educational or medical reasons.

Where such intent is clear, we make allowances for the content. For example, while we restrict some images of female breasts that include the nipple, we allow other images, including those depicting acts of protest, women actively engaged in breast-feeding and photos of post-mastectomy scarring. For images depicting visible genitalia or the anus in the context of birth and after-birth moments or health-related situations we include a warning label so that people are aware that the content may be sensitive. We also allow photographs of paintings, sculptures, and other art that depicts nude figures.

Do not post:

Images of

- Real nude adults, where nudity is defined as
    - Visible genitalia except in the context of birth giving and after-birth moments or health-related situations (for example, gender confirmation surgery, examination for cancer or disease prevention/assessment)
    - Visible anus and/or fully nude close-ups of buttocks unless photoshopped on a public figure
    - Uncovered female nipples except in the context of breastfeeding, birth giving and after-birth moments, health-related situations (for example, post-mastectomy, breast cancer awareness or gender confirmation surgery) or an act of protest
- Sexual activity, including
    - Sexual intercourse
        - Explicit sexual intercourse, defined as mouth or genitals entering or in contact with another person's genitals or anus, where at least one person's genitals are nude
        - Implied sexual intercourse, defined as mouth or genitals entering or in contact with another person's genitals or anus, even when the contact is not directly visible, except in cases of a sexual health context, advertisements, and recognized fictional images or with indicators of fiction
        - Implied stimulation of genitalia/anus, defined as stimulating genitalia/anus or inserting objects

into genitalia/anus, even when the activity is not directly visible, except in cases of sexual health context, advertisements, and recognized fictional images or with indicators of fiction
- Other sexual activities including (but not limited to)
  - Erections
  - Presence of by-products of sexual activity
  - Stimulating genitals or anus, even if above or under clothing
  - Use of sex toys, even if above or under clothing
  - Stimulation of naked human nipples
  - Squeezing female breasts, defined as a grabbing motion with curved fingers that shows both marks and clear shape change of the breasts. We allow squeezing in breastfeeding contexts
- Fetish content that involves
  - Acts that are likely to lead to the death of a person or animal
  - Dismemberment
  - Cannibalism
  - Feces, urine, spit, snot, menstruation, or vomit
- Digital Content that meets our definition of sexual activity, except when posted in an educational or scientific context, or when it meets one of the criteria below and shown only to individuals 18 years and older.

We only show this content to individuals 18 and older:

- Real world art that depicts sexual activity
- Posting photographs or videos of objects that depict sexual activity in real world art
- Implied sexual activity in advertisements
- Implied sexual activity in recognized fictional images or with indicators of fiction
- Digital content that meets our definition of sexual activity, where:
  - The sexual activity (intercourse or other sexual activities) isn't directly visible
  - Content was posted in a satirical or humorous context
  - Only body shapes or contours are visible

15. Sexual Solicitation

III. Objectionable Content

Policy Rationale

As noted in Section 8 of our Community Standards (Sexual Exploitation of Adults), people use Facebook to discuss and draw attention to sexual violence and

exploitation. We recognize the importance of and want to allow for this discussion. We draw the line, however, when content facilitates, encourages or coordinates sexual encounters or commercial sexual services between adults such as prostitution or escort services. We do this to avoid facilitating transactions that may involve trafficking, coercion, and non-consensual sexual acts. We also restrict sexually explicit language that may lead to solicitation because some audiences within our global community may be sensitive to this type of content and it may impede the ability for people to connect with their friends and the broader community.

Do not post:

Attempting to coordinate adult commercial sexual services or prostitution activities, such as requesting or offering or asking for rates for escort services and paid sexual fetish or domination services.

Attempted coordination of or recruitment for adult sexual activities, including but not limited to:

> Filmed sexual activities
> Pornographic activities, strip club shows, live sex performances, erotic dances
> Sexual, erotic, or tantric massages

Explicit sexual solicitation by, including but not limited to the following, offering or asking for:

> Sex or sexual partners
> Sex chat or conversations
> Nude photos/videos/imagery

Content that is implicitly or indirectly offering or asking for sexual solicitation and meets both of the following criteria. If both criteria are not met, it is not deemed to be violating. For example, if content is a hand-drawn image depicting sexual activity but does not ask or offer sexual solicitation, it is not violating:

> Criteria 1: Offer or Ask
>
>> Content implicitly or indirectly[*] (typically through providing a method of contact) offers or asks for:
>>
>>> Nude imagery, or
>>>
>>> Sex or sexual partners, or
>>>
>>> Sex chat conversations
>
> Criteria 2: Suggestive Elements
>
>> Content makes the aforementioned offer or ask using one of the following sexually suggestive elements:

> > > Contextually specific and commonly sexual emojis or emoji strings, or
>
> > > Regional sexualized slang, or
>
> > > Mentions or depictions of sexual activity (including hand drawn, digital, or real world art) such as: sexual roles, sex positions, fetish scenarios, state of arousal, act of sexual intercourse or activity (sexual penetration or self-pleasuring), or
>
> > > Imagery of real individuals with nudity covered by human parts, objects, or digital obstruction, including long shots of fully nude butts

An offer or ask for pornographic material (including, but not limited to, sharing of links to external pornographic websites)

Sexually explicit language that goes into graphic detail beyond mere reference to:

> A state of sexual arousal (wetness or erection) or
> An act of sexual intercourse (sexual penetration, self-pleasuring or exercising fetish scenarios)

16. Cruel and Insensitive

III. Objectionable Content

Policy Rationale

We believe that people share and connect more freely when they do not feel targeted based on their vulnerabilities. As such, we have higher expectations for content that we call cruel and insensitive, which we define as content that targets victims of serious physical or emotional harm.

We remove explicit attempts to mock victims and mark as cruel implicit attempts, many of which take the form of memes and GIFs.

Do not post:

Content that depicts real people and laughs at or makes fun of their serious physical injury, starvation, or serious or fatal disease or disability.Content that contains sadistic remarks and any visual or written depiction of real people experiencing premature death, serious physical injury, physical violence or domestic violence.

Imagery that depicts real animals that are visibly experiencing and being laughed at, made fun of, or contain sadistic remarks for any of the following (except staged animal vs. animal fights or animal fights in the wild):

premature death
serious physical injury (including mutilation)
physical violence from a human

17. Account Integrity and Authentic Identity

IV. Integrity and Authenticity

Policy Rationale

Authenticity is the cornerstone of our community. We believe that authenticity helps create a community where people are to each other, and to Facebook, in meaningful ways. We want to allow for the range of diverse ways that identity is expressed across our global community, while also preventing impersonation and identity misrepresentation. That's why we require people to connect on Facebook using the name they go by in everyday life. Our authenticity policies are intended to create a safe environment where people can trust and hold one another accountable.

In order to maintain a safe environment and empower free expression, we remove accounts that are harmful to the community, including those that compromise the security of other accounts and our services. We have built a combination of automated and manual systems to block and remove accounts that are used to persistently or egregiously abuse our Community Standards.

Because account level removal is a severe action, whenever possible, we aim to give our community a chance to learn our rules and follow our Community Standards. Penalties, including account disables, are designed to be proportionate to the severity of the violation and the risk of harm posed to the community. Continued violations, despite repeated warnings and restrictions, or violations that pose severe safety risks will lead to an account being disabled.

We do not allow the use of our services and will disable accounts if you:

> <u>Severely violate</u> our Community Standards.
> <u>Persistently violate</u> our Community Standards.
> Represent <u>Dangerous Individuals or Organizations</u>
> Create or use an account that demonstrates an intent to violate our Community Standards.
> Create an account, page, group or event to evade our enforcement actions, including creating an account to bypass a restriction or after we've disabled your previous account, page, group or event.
> Impersonate others by :
>> Using their photos with the explicit aim to deceive others.
>>
>> Creating an account assuming to be or speak for another person or entity.

Creating a Page assuming to be or speak for another person or entity for whom the user is not authorized to do so.

Are under 13 years old.

Are a convicted sex offender.

Are prohibited from receiving our products, services or software under applicable laws.

In certain cases, we will seek further information about an account before taking actions ranging from temporarily restricting accounts to permanently disabling them.

Accounts misrepresenting their identity (Facebook and Messenger only) by:

Using a name that does not abide by our name policies.

Providing a false date of birth.

Creating a single account that represents or is used by more than one person.

Maintaining multiple accounts as a single user.

Compromised accounts.

Empty accounts with prolonged dormancy.

18. Spam

IV. Integrity and Authenticity

Policy Rationale

We work hard to limit the spread of spam because we do not want to allow content that is designed to deceive, or that attempts to mislead users to increase viewership. This content creates a negative user experience and detracts from people's ability to engage authentically in online communities. We also aim to prevent people from abusing our platform, products, or features to artificially increase viewership or distribute content en masse for commercial gain.

Do not:

Post, share, engage with content, or create accounts, groups, pages, events or other assets, either manually or automatically, at very high frequencies. Attempt to or successfully sell, buy, or exchange site privileges or product features, such as accounts, admin roles, permission to post, pages, groups, likes, etc., except in the case of clearly identified branded content, as defined by our Branded Content Policy.

Require or claim that users are required to engage with content (e.g. liking, sharing) before they are able to view or interact with promised content.

Encourage likes, shares, follows, clicks, or the use of apps or websites under false pretenses, such as:

Offering false or non-existent services or functionality (e.g., "Get a 'Dislike' button!")

Failing to direct to promised content (e.g., "Click here for a discount code at Nordstrom"; false play buttons)

Directing users off of Facebook through the deceptive or misleading use of URLs, defined as:

Cloaking: Presenting different content to Facebook users and Facebook crawlers or tools.

Misleading content: Content contains a link that promises one type of content but delivers something substantially different. Deceptive Pop-up websites: Websites that require an action (captcha, watch ad, click here) in order to view the expected landing page content and the domain name of the URL changes after the required action is complete.
Like / Share-Gating: Landing pages that require users to like, share, or otherwise engage with content before gaining access to off-Facebook content.

Deceptive Landing Page Functionality: Websites that have a misleading user interface, which results in accidental traffic being generated (pop-ups/unders, clickjacking, etc.)

Impersonation: The website pretends to be a reputable brand or service by usin   g a name, domain, or content featuring typos, misspellings, or other means to impersonate well-known brands (e.g. www.faceb00k.com, www.face_book.com, using a landing page similar to another, trusted site to mislead visitors).

And other behaviors that are substantially similar to those listed above.

19. Cybersecurity

IV. Integrity and Authenticity

Policy Rationale

We recognize that the safety of our users includes the security of their personal information and accounts, profiles, and other Facebook entities they may manage, as well as our products and services more broadly. Attempts to gather sensitive personal information or engage in unauthorized access by deceptive or invasive methods are harmful to the authentic, open, and safe atmosphere that we want to

foster. Therefore, we do not allow attempts to gather sensitive user information through or engage in unauthorized accessthe abuse of our platform and products, or services.

Do not:

Attempt to compromise user accounts, profiles, or other Facebook entities, abuse our products or services, gather sensitive information through deceptive means, or attempt to engage in unauthorized access, including:

- Gaining access to accounts , profiles, Facebook, Inc. entities, or user data other than your ownthrough deceptive means or without explicit permission from the account, profile, or entity owner.
- Encouraging or deceiving users to download or run files or programs that will compromise a user's online or data security, including through malicious software or websites.. Such files and programs will be deemed malicious software or "malware" if they harm or gain unauthorized access to a computer, device, or network.
- Attempting to obtain, acquire or request another user's login credentialsor other sensitive information, , whether explicitly or through deceptive means such as phishing (e.g. fake surveys designed to capture log-in info or links to fake login pages or impostor websites)or the use of malicious software or websites..
- Publicly sharing your own or others' login information, either on platform or through a third party service.
- Creating, sharing, or hosting malicious software browser extensions and mobile applications, on or off platform that put our users or products and services at risk.
- Providing online infrastructure, including Web Hosting Services, Domain Name System servers, and ad networks that enables abusive links such that a majority of those links on Facebook or Instagram violate the Spam or Cybersecurity sections of the Community Standards.

20. Inauthentic Behavior

IV. Integrity and Authenticity

Policy Rationale

In line with our commitment to authenticity, we don't allow people to misrepresent themselves on Facebook, use fake accounts, artificially boost the popularity of content, or engage in behaviors designed to enable other violations under our Community Standards. This policy is intended to create a space where people can trust the people and communities they interact with.

Do not:

Use multiple Facebook accounts or share accounts between multiple people

Misuse Facebook or Instagram reporting systems to harass others

Conceal a Page's purpose by misleading users about the ownership or control of that Page

Engage in or claim to engage in Inauthentic Behavior, which is defined as the use of Facebook or Instagram assets (accounts, pages, groups, or events), to mislead people or Facebook:

> about the identity, purpose, or origin of the entity that they represent

> about the popularity of Facebook or Instagram content or assets

> about the purpose of an audience or community

> about the source or origin of content

> to evade enforcement under our Community Standards

Engage in, or claim to engage in Coordinated Inauthentic Behavior, defined as the use of multiple Facebook or Instagram assets, working in concert to engage in Inauthentic Behavior (as defined above), where the use of fake accounts is central to the operation

Engage in or claim to engage in Foreign or Government Interference, which is Coordinated Inauthentic Behavior conducted on behalf of a foreign or government actor.

21. False news

IV. Integrity and Authenticity

Policy rationale

Reducing the spread of false news on Facebook is a responsibility that we take seriously. We also recognise that this is a challenging and sensitive issue. We want to help people stay informed without stifling productive public discourse. There is also a fine line between false news and satire or opinion. For these reasons, we don't remove false news from Facebook, but instead significantly reduce its distribution by showing it lower in the News Feed. Learn more about our work to reduce the spread of false news <u>here</u>.

We are working to build a more informed community and reduce the spread of false news in a number of different ways, namely by

> Disrupting economic incentives for people, Pages and domains that propagate misinformation
> Using various signals, including feedback from our community, to inform a machine learning model that predicts which stories may be false

Reducing the distribution of content rated as false by independent third-party fact-checkers

Empowering people to decide for themselves what to read, trust and share by informing them with more context and promoting news literacy

Collaborating with academics and other organisations to help solve this challenging issue

22. Manipulated Media

IV. Integrity and Authenticity

Policy Rationale

Media, including image, audio, or video, can be edited in a variety of ways. In many cases, these changes are benign, like a filter effect on a photo. In other cases, the manipulation isn't apparent and could mislead, particularly in the case of video content. We aim to remove this category of manipulated media when the criteria laid out below have been met.

In addition, we will continue to invest in partnerships (including with journalists, academics and independent fact-checkers) to help us reduce the distribution of false news and misinformation, as well as to better inform people about the content they encounter online.

For the following Community Standards, we require additional information and/or context to enforce:

Do not post:

Videos that have been edited or synthesized, beyond adjustments for clarity or quality, in ways that are not apparent to an average person, and would likely mislead an average person to believe:

A subject of the video said words that they did not say, AND

The video is the product of artificial intelligence or machine learning, including deep learning techniques (e.g., a technical deepfake), that merges, combines, replaces, and/or superimposes content onto a video, creating a video that appears authentic.

This policy does not extend to content that is parody or satire or is edited to omit words that were said or change the order of words that were said.

23. Memorialization

IV. Integrity and Authenticity

Policy Rationale

When someone passes away, friends and family can request that we memorialize the Facebook account. Once memorialized, the word "Remembering" appears above the name on the person's profile to help make it that the account is now a memorial site and protects against attempted logins and fraudulent activity. To respect the choices someone made while alive, we aim to preserve their account after they pass away. We have also made it possible for people to identify a legacy contact to look after their account after they pass away. To support the bereaved, in some instances we may remove or change certain content when the legacy contact or family members request it.

Visit Hard Questions for more information about our memorialization policy and process. And see our Newsroom for the latest tools we're building to support people during these difficult times.

For victims of murder and suicide we will remove the following content if it appears on the deceased's profile photo, cover photo, or among recent timeline posts when requested by a legacy contact or family member of the deceased:

> Visual depiction of the object used in the deceased's death
> Imagery of the convicted or alleged murderer of the deceased
> Content related to the deceased's death

For victims of murder, we will also remove the convicted or alleged murderer from the deceased's profile if referenced in relationship status or among friends.

For the following Community Standards, we require additional information and/or context to enforce:

When brought to our attention by the legacy contact or a family member of the deceased, we will:

> Remove violating comments on a memorialized profile, which would typically require the individual to self report so that we know that they are unwanted
> Remove praise or support for the death, disease, or harm of the deceased person on a memorialized profile
> Change the deceased's individual's privacy settings from public to friends-only when there is harmful content on the profile
> Change a violating account name on the profile of the deceased individual

24. Intellectual Property

V. Respecting Intellectual Property

Policy rationale

Facebook takes intellectual property rights seriously and believes that they are important to promoting expression, creativity and innovation in our community. You own all of the content and information that you post on Facebook, and you control how it is shared through your privacy and application settings. However, before sharing content on Facebook, please make sure that you have the right to do so. We ask that you respect other people's copyrights, trademarks and other legal rights. We are committed to helping people and organisations promote and protect their intellectual property rights. Facebook's Terms of Service do not allow people to post content that violates someone else's intellectual property rights, including copyright and trademark. We publish information about the intellectual property reports that we receive in our bi-annual Transparency Report, which can be accessed at https://transparency.facebook.com/

Upon receipt of a report from a rights holder or an authorised representative, we will remove or restrict content that engages in:

> Copyright infringement
> Trademark infringement

25. User requests

VI. Content-Related Requests and Decisions

We comply with:

> User requests for removal of their own account
> Requests for removal of a deceased user's account from a verified immediate family member or executor
> Requests for removal of an incapacitated user's account from an authorised representative

26. Additional Protection of Minors

VI. Content-Related Requests and Decisions

We comply with:

> Requests for removal of an underage account
> Government requests for removal of child abuse imagery depicting, for example, beating by an adult or strangling or suffocating by an adult
> Legal guardian requests for removal of attacks on unintentionally famous minors

For the following Community Standards, we require additional information and/or context to enforce:

We may remove content created for the purpose of identifying a private minor if there may be a risk to the minor's safety when requested