

GLOSSARY OF TERMS

1737 / Need to Talk?

New Zealand's free 24/7 mental-health support line. Call or text 1737 to speak with a trained counsellor.

Accessibility (Digital)

Designing online content so it's usable by people with disabilities — for example, alt text, captions, and clear layout.

Admin Assist (Groups)

A Facebook Group automation tool that pre-approves, declines, or flags posts and members based on rules you set.

AI (Artificial Intelligence)

Technology that performs human-like tasks such as filtering spam or summarising comments; used behind many Meta moderation tools.

Approved Agency

Under the Harmful Digital Communications Act 2015, Netsafe is the Approved Agency responsible for receiving complaints, assisting victims, and working with online content hosts to resolve harm.

Burnout

Emotional and mental exhaustion caused by prolonged stress; common among moderators exposed to harmful or high-volume content.

Communication Principles

Ten behavioural principles under the HDCA that outline what counts as harmful digital communication (e.g. no harassment, no false allegations, no incitement).

Community Standards

Platform rules on acceptable behaviour. Breaches can lead to post removal or account restrictions even if NZ law isn't broken.

Content Warning (CW)

A note at the start of a post alerting readers to sensitive or distressing content (e.g. violence, trauma, abuse). Sometimes know as Trigger Warning (TW).

Coroners Act 2006, Section 71 (Reporting Suicide)

Restricts publication of a suspected suicide's method or details until the Chief Coroner releases findings.

Crisis Terms

Words or phrases that signal distress or imminent risk and should prompt escalation.

Cultural Safety

Ensuring online spaces respect and protect the identities and wellbeing of people from all cultural backgrounds.

De-escalation

Moderator steps that reduce tension, using calm language, slowing comments, or temporarily pausing discussion.

Deepfakes

AI-generated or altered video, images, or audio that convincingly depict people doing or saying things they never did.

Defamation

Publishing false information that harms someone's reputation; truth and honest opinion are defences.

DIA – Department of Internal Affairs

NZ government agency responsible for addressing objectionable or illegal online content through its Digital Safety Group.

Digital Communication

Any form of online or electronic messaging, including emails, social media posts, comments, images, and videos.

Dog-whistle

Coded language that seems harmless publicly but signals bias or hostility to a targeted group.

Doxxing

Publishing or sharing someone’s private or identifying information (address, workplace, phone number etc.) without consent.

Escalation

Referring a serious or unlawful issue (e.g. threat, harassment, exploitation) to Netsafe, Police, the DIA, or another agency.

Featured Post

A Facebook post pinned to the top of a group feed to highlight important or time-sensitive information.

FVPC Act (Films, Videos and Publications Classification Act 1993)

NZ law defining and regulating “objectionable” content such as child sexual exploitation, extreme violence, or terrorism material.

Guides (Groups)

A Facebook feature allowing key posts to be grouped into themes (Welcome, Safety, FAQs) for easy reference.

Harmful Digital Communications Act 2015 (HDCA)

NZ law to deter, prevent, and mitigate online harm and provide quick redress for victims of digital communications causing serious emotional distress.

Holding Post

A temporary pinned notice used during high tension to pause discussion and explain next steps.

Holding Statement

A neutral message to pause or de-escalate conflict (e.g. “We’re reviewing this post, please pause kōrero for now.”).

Kaupapa

A Māori term for purpose or foundational principle. In moderation, it anchors a group’s values and decisions.

Kaupapa Māori

Māori-led frameworks and values that guide how communities are built and moderated.

Keyword Alerts (Groups)

Facebook tool that notifies moderators when chosen words or phrases appear in posts or comments.

Kōrero

Māori for “conversation” or “discussion.” Used to describe respectful dialogue within online communities.

Live Stream

A real-time video broadcast. Moderators should monitor streams closely to prevent harmful or illegal content.

Mana

Authority, dignity, and prestige. Moderators maintain mana by acting fairly and respectfully.

Manaakitanga

A Māori value meaning care, hospitality, and respect — hosting others kindly and safeguarding community wellbeing.

Misinformation / Disinformation

False or misleading information shared without (misinformation) or with (disinformation) intent to deceive.

Mod Chat

A private channel where moderators coordinate actions and support each other.

Mod Team / Co-Moderators

The group of people sharing moderation duties and decision-making for a community.

Moderator (Mod)

A person who manages an online community, enforces rules, and keeps discussions safe and respectful.

Moderation Assist (Pages)

Meta's automated tool for public Pages that hides or reviews comments based on filters such as keywords or spam.

Moderation Log (Mod Log)

A private record of moderation actions — what was removed, why, and by whom.

Multicultural Safety

Ensuring online spaces respect and protect the identities and wellbeing of people from all cultures, including Māori and Pacific communities.

Objectionable Content

Material depicting sex, crime, cruelty, or violence that's injurious to the public good. Includes child sexual exploitation, bestiality, and terrorism imagery.

Online Content Host

Anyone who owns or controls an online space and can moderate or delete content within it.

Peer Support

Emotional or practical support between moderators to manage stress and prevent burnout.

Pile-on

When multiple users attack or criticise one person in a thread, often escalating into harassment.

Post Approval

A Facebook Group setting requiring moderator review before posts become visible.

Privacy Act 2020

NZ law governing how personal information is collected, used, and shared.

Psychological Safety / Safe Space

An environment where people can participate without fear of ridicule, harassment, or retaliation.

Scam / Fraud

A deceptive scheme designed to steal money, data, or personal information.

Serious Emotional Distress

The threshold of harm under the HDCA – sustained fear, humiliation, or distress beyond ordinary offence.

Tap-Out System

A wellbeing practice allowing moderators to step back temporarily when overwhelmed.

Tangata Whenua

Māori as the Indigenous people of Aotearoa New Zealand.

Te ao Māori

The Māori worldview, values, language, and customs shaping understanding and behaviour.

Te reo Māori

The Māori language, an official language of New Zealand and encouraged in online spaces.

Tikanga

Māori customs and protocols that guide behaviour and decision-making in culturally safe spaces.

Tone Policing

Criticising how someone expresses a point instead of addressing the issue itself, which can silence marginalised voices.

Tone Reset

A moderator comment that re-centres respectful dialogue and reminds members of the group's kaupapa.

Trauma-Aware Moderation

Recognising that people carry past harm and moderating with empathy and non-blaming language.

Values Statement

A short summary of shared principles (e.g. empathy, inclusivity, safety) that guide moderator decisions.

Wellbeing Check-In / Debrief

A structured chat after handling harmful content to reflect and support each other.

Whakapapa

Māori term for genealogy or lineage, the relationships that define identity and belonging.

Whanaungatanga

A Māori concept of connection, relationship, and shared belonging, the foundation for inclusive moderation.